DOCUMENT RESUME

ED 109 889 FL 006 644

AUTHOR Kohonen, Viljo, Ed.
TITLE Inter-University Student Selection for English 1974: A Joint Entrance Test Project of the Universities of Joensuu, Jyvaskyla and Turku.
INSTITUTION Jyvaskyla Univ. (Finland).
REPORT NO ISBN-951-677-388-5
PUB DATE 74
NOTE 83p.

EDRS PRICE MF-$0.76 HC-$4.43 PLUS POSTAGE
DESCRIPTORS *Admission Criteria; Competitive Selection; *English (Second Language); *Higher Education; *Language Tests; Statistical Analysis; Student Attitudes; Student Evaluation; *Test Construction; Test Results
IDENTIFIERS *Finland

ABSTRACT

The report gives the results of the 1974 joint student selection entrance test for English, administered simultaneously at the universities of Joensuu, Jyvaskyla and Turku. The test battery comprised 309 multiple-choice items and was completed by 843 candidates. Hirvonen's paper examines problems of student selection and theoretical aspects of language testing. Kohonen gives a brief account of the planning and administration of the examination and provides the main results of a statistical analysis of the tests. Jarrett examines the translation analysis test, and Hughes and May analyze the listening tests. Roininen's paper is a sociological study of the candidates' opinions about the selection procedures and the tests. In each paper, some suggestions have been offered for the further development of student selection. The research was carried out under a grant from the Ministry of Education. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

KIELIKESKUKSEN JULKAISUJA
REPORTS FROM THE LANGUAGE CENTRE
JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

ED107889

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
University of Jyväskylä
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

FL006644

INTER-UNIVERSITY STUDENT SELECTION FOR ENGLISH 1974: A JOINT ENTRANCE TEST PROJECT OF THE UNIVERSITIES OF JOENSUU, JYVÄSKYLÄ AND TURKU

5/1974

Edited by Viljo Kohonen

INTER-UNIVERSITY STUDENT SELECTION FOR ENGLISH 1974: A JOINT ENTRANCE TEST PROJECT OF THE UNIVERSITIES OF JOENSUU, JYVÄSKYLÄ AND TURKU

Edited by Viljo Kohonen

CONTENTS

PREFACE

1. Pekka Hirvonen: PROBLEMS OF UNIVERSITY STUDENT SELECTION WITH SPECIAL REFERENCE TO ENGLISH DEPARTMENTS 1

2. Viljo Kohonen: JOINT ENTRANCE TEST JOENSUU, JYVÄSKYLÄ, TURKU, 1974: SELECTION PROCEDURES AND TEST ANALYSIS 18

3. Peter Jarrett: OBSERVATIONS ON THE 1974 TRANSLATION ANALYSIS TEST 30

4. Glyn Hughes and Anthony May: ASPECTS OF THE 1974 LISTENING TESTS 33

5. Väinö Roininen and Pirkko Kosonen: THE 1974 JOINT STUDENT SELECTION FOR ENGLISH: CANDIDATES' OPINIONS ABOUT THE INFORMATION, THE TEST BATTERY USED AND THE INTER-UNIVERSITY CO-OPERATION 49

APPENDIX 1. Advance information for candidates applying to Turku (from Opintielle Turun yliopistoon 1974). 62

APPENDIX 2. Test instructions. 64

APPENDIX 3. Questionnaire. 65

APPENDIX 4. Advance information of the university of Jyväskylä (from Jyväskylän yliopiston opinnot ja opiskelijavalinta v. 1974). . . 69

APPENDIX 5. Factor analysis of some background variables and opinions on the appropriateness of the different tests. 70

APPENDIX 6. Footnote to page 36. 71

PREFACE

The present publication gives the main results of a joint student selection project for English, carried out at the universities of Joensuu, Jyväskylä and Turku in summer 1974. The cooperation dates back to the first annual joint conference of the English departments of the universities of Finland, held at Lammi in November 1972, where the idea of inter-departmental student selection was conceived with the aim of sharing the work load entailed in the preparation and administration of a comprehensive language test battery. The proposed project subsequently came to include the universities of Joensuu, Jyväskylä and Turku, Viljo Kohonen becoming the project leader. The first joint entrance test of these three departments was administered in summer 1973, and the cooperation was continued in 1974, encouraged by the positive experiences from the previous summer.

In order to disseminate the results and ideas to a wider readership for discussion and criticism, the planning team turned to the Ministry of Education for a research grant for the 1974 joint student selection, and a grant was put at the research group's disposal on May 16th, 1974. The jointly planned entrance test was administered simultaneously at the three universities on July 8th.

The present publication consists of 5 papers written by members of the planning team, covering some of the central problems and results. As inter-departmental cooperation is currently a topical question at the language departments of the Finnish universities, it is hoped that these papers will benefit others by clarifying some of the problems and offering suggestions for solving them. It should be emphasized that the present test battery developed by the planning team does not pretend to be "the entrance test"; it is just one manageable solution, which needs to be developed and modified in the light of the changing aims of university studies and, consequently, the changing requirements with respect to the student characteristics desired.

Jyväskylä, November 1st, 1974.

Viljo Kohonen

Pekka Hirvonen:

PROBLEMS OF UNIVERSITY STUDENT SELECTION, WITH SPECIAL REFERENCE TO ENGLISH DEPARTMENTS

00. PROLOGUE

Too little discussion has been going on in our English departments on the issues involved here, especially on the aims of instruction, which are a necessary background to a discussion of student selection criteria and procedures.

The main purpose of this article is to provoke discussion. The article is argumentative, and to remind the reader of this I have replaced the customary, respectable "in view of these results the present writer is inclined to conclude that" with "I think", which is what the former phrase means anyway. I believe the issues involved here should be thrashed out in the open. If this article manages to elicit either opposing or supporting views and arguments from my learned colleagues and if in the ensuing interchange some issues are clarified or settled, then this article will not have been a total waste of our ever-diminishing supply of paper.

0. INTRODUCTION

The essential problem in university student selection seems to be in determining what qualities to look for in the applicants and then finding ways of assessing those qualities. In the Finnish situation, discriminating only between the qualified ones and the unqualified ones is normally not enough: because the number of qualified applicants is usually more than can be admitted, our selection instruments must discriminate <u>among</u> the qualified ones as well.

It is, however, rather difficult to determine what qualities to look for in the applicants unless the objectives of their university studies have been agreed on: what are they to learn, and to what end?

Traditionally, our departments of English (or the departments of English Philology, to be more accurate) have had no explicit objectives of teaching (the 'development of critical thinking', sometimes offered as an answer, is not relevant in this discussion because that 'aim' must be taken for granted in <u>all</u> university education worth the name). Indeed, some of our English Philologists react to the words "objective", "goal", or "aim" as if

they were swearwords (personal observation at a recent joint conference of departments of English and English Philology).

Let me therefore begin my discussion of student selection with a tentative discussion of the objectives of university English teaching.

1. ON THE OBJECTIVES OF UNIVERSITY ENGLISH TEACHING

For a long time, the great majority of those who have studied English Philology have eventually become teachers of English at the primary or secondary level; a degree in English Philology is one of the official qualifications for teaching posts. These known facts have had little influence on the contents and methods of study. The general aim (not expressed but implicit in the syllabuses) seems to have been to train English Philologists, i.e. people who could take up philological research - study of old (classical, literary) texts - and as university teachers pass on their knowledge of English Philology to new generations of students. With English Philology offering career opportunities to less (probably far less) than five per cent of the graduates, it seems that the interests and needs of the great majority have been sacrificed in favour of a small minority.

Recently, however, some clear indications of change have appeared. At the end of 1969, the Ministry of Education appointed a committee for the reform of studies and degrees in the faculties of the humanities and the social sciences. This committee turned in its report in 1972 ("FYTT": Committee Report 1972: A 17). After a lengthy nationwide discussion of the report's principles and proposals, the Ministry of Education in the spring of 1974 took a firm stand in favour of the main proposals in the "FYTT" report concerning the qualitative development of university instruction (Opetusministeriön kannanotto 13.3.1974 korkeakoulujen tutkinnonuudistuksen jatkosuunnittelusta). Among the central principles, now to be regarded as official policy, are the following:

1. Problem-centeredness. The general starting-point of instruction and the basis of the selection and systematization of the subject matter must be in the real problems existent in society.
2. Interaction between teaching/studying and research must be improved and extended from the present state.
3. Interdisciplinarity. Problem-centeredness presupposes that the matters to be dealt with in the teaching and studying can be examined from the points of view of several different disciplines.

4. Profession-orientedness. In the Humanities and the Social Sciences in particular, there is a clear need for the studies to be more directly oriented towards some scientific, social, or professional activities.

5. Development of general abilities. In university training, general scientific and methodological abilities should be emphasized. Instead of narrow specialization, the training should develop a facility for acquiring specific professional abilities and skills and for applying the acquired knowledge.

Let us examine points 1 and 4 above closely. What are the problems and the professions in our society towards which university English studies should be oriented? A moment's reflection will reveal that the largest problem area is the planning, carrying out, and development-through-research of the teaching of English as a foreign language and the production of learning materials. Other areas include the linguistic study and description of the structure and use of English, both as an end in itself and particularly as a theoretical basis for the first-mentioned activities. Translation and dealing with English-medium materials in the media, foreign service and business also seem to require university graduates in English.

Another factor of great importance for the goal-setting is the Teacher Training Act (passed in 1972), which came into force in the autumn of 1974 (at the universities of Joensuu and Jyväskylä, a year earlier). This Act puts the universities in charge of all teacher training for primary and secondary education, and thus, for example, assigns to the departments of English the responsibility for the 'subject-matter' training of English teachers. From now on, remarks like "If some or most of our graduates end up teaching the language at school, that's none of our business" are no longer available as excuses for outdated, irrelevant, or inefficient teaching.

Clearly, future teachers of English are going to continue forming the majority among our customers; we are, moreover, required by law to provide for the needs of teachers and teaching of English. Does this mean that we should turn our departments into high-level vocational schools and concentrate solely on methodological know-how, with the aim of producing highly specialized teachers ready to enter the classroom? - In my view, we cannot and we must not. In the first place, the training of English teachers is the joint responsibility of the departments of English, the departments of Education, and the training schools. The function of the English department in this co-ordinated effort must be seen as that of giving the necessary linguistic background that enables the student to acquire, critically evaluate and develop the methodologies and techniques that he comes

across in the specialized training, in the literature, and in actual teaching. In the second place, the basic degree in university studies (i.e. "fil. kand.") cannot be tailored to the needs of one specific profession only (cf. principle 5 on page 3): it must provide a sound theoretical and skill basis for several related professions, with specialization coming mostly afterwards in the form of further studies (e.g. didactic studies and practice teaching for teachers, "FL" and "FT" degrees for researchers) or practical training on the job, or both. This basis must be equally valid for those aiming at teaching, research, translation, or high-level English-medium work in the media or foreign service or with publishing houses (I doubt whether this list exhausts the possibilities). Separate training programmes for these professions would be highly undesirable because the students' future careers cannot be predicted reliably. They are subject to change both before and after graduation, and a change of professions within the same broad field should always be possible with a reasonable amount of further study or training.

It seems then that university English studies up to the basic degree should strongly concentrate on those areas and abilities that are essential for all high-level professional work pertaining to the English language. Tentatively, I suggest that these areas might be the following:

1. A firm practical command of the language: ability to understand others and express himself confidently, accurately and fluently in both spoken and written English, including a practical command of the main varieties of English according to region, social standing, and attitude.

2. Familiarity with the language background. For practical reasons, this may have to be restricted mostly to Britain and the USA: everyday life, generally held attitudes, social, economic and political conditions, institutions, "high" culture (e.g. literature), foreign policies; the role of English as a world language and its consequences.

3.1. Knowledge of general linguistics: the nature and structure of human language; how it changes; how it is acquired and how it is used in different situations; language universals and how individual languages differ from each other; essentials of research methodology.

3.2. Theoretical knowledge of English: the structure of modern English from the viewpoint of all levels of linguistic analysis; varieties according to region, social standing, subject matter, attitude, medium;

the salient features of the historical development of English.

4. Ability and willingness to apply the above-mentioned knowledge and skills to the linguistic problems in society.

Proper coverage of these areas and the acquisition of these skills would certainly take up most of the study time up to the level of 'laudatur', but some specialization could take place towards the end of these studies in the form of elective courses of study, such as further linguistic (psycholinguistic, sociolinguistic, historical) studies of English, contrastive analysis and error analysis, topics in applied linguistics - theoretical bases for language teaching (goal-setting, methodology, evaluation), theory and practice of translation, or English literature. Moreover, the choice of other subjects provides, of course, plenty of scope for specialization even at the undergraduate level: Linguistics, Phonetics, Native Language, Literature, Education, Psychology, Sociology ...

The above outline presents the guiding principles on which the syllabus of English studies at the University of Joensuu is based. As it is not altogether different from what some of our other English departments seem to be moving towards (judging by their recent "study requirements" (='tutkintovaatimukset') and as there are no explicit alternative sets of general aims available, I shall in the following discussion use the above outline as the background against which desirable student qualities in the selection process may be determined.

2. ON DESIRABLE STUDENT QUALITIES

Like all other university departments, the department of English should expect students with an ability not only to memorize and recall knowledge acquired from textbooks and lectures but to apply it to solving problems in their field. This is, of course, an ability that can, and should, be acquired in the course of university studies, and it certainly can be taught if its importance as a general aim is recognised. But what is the underlying aptitude that enables a student to acquire this ability, and how can it be measured? Apparently some evidence of the applicants' ability to handle theoretical information is obtainable in their school records and the marks in their matriculation examination - which indeed are used as selection criteria by many departments. What disturbs me in the use of these criteria is that a pupil can, as we know, get quite far at school (and all too often at university as well) by hard work at

amassing and memorizing details and correctly repeating them in a test situation. It seems not at all certain that a student selected on these criteria will turn out to be the best possible university student and - more importantly - a successful practising English linguist.

In university study of a foreign language, motivation is important: interest in the language and desire to know more about its structure, its functioning, its background, its development. Again, these qualities are difficult to measure directly with our present knowledge and resources, but indirect evidence may be furnished by the student's past success in language studies.

University English studies clearly require more than average "verbal reasoning ability", or "verbal intelligence", or "language/linguistic aptitude" (or what should I call it). By this I mean that while most people can, in favourable conditions, probably learn to communicate tolerably well in a foreign language, the ability to analyse and describe the language in rigorous theoretical terms - an essential aspect of university study - requires more, and perhaps a different kind of, aptitude. Even though indirect evidence of the student's verbal reasoning ability and his sensitivity to language can be had from his past language achievement, I think there is a need for more direct assessment of these qualities, which I consider equally relevant for research into the language and the teaching of it. Perhaps we should take a good look at the existing tests of verbal reasoning ability with a view to their possible applicability for student selection.

A good practical command of the language, both oral and written, is obviously essential. It seems to me that it is precisely because of the deficiencies of the existing measures of the applicants' language skills - the school record and the matriculation examination in English - that university English departments have started using entrance tests of their own. One might argue, of course, that the language skills should not be the main criterion of selection because these skills can be taught at the university, and in many cases more efficiently than at school. But considering that nearly all the textbooks are in English and that English is extensively used as the medium of instruction from the very beginning, one must conclude that the student's initial language skills are a very important factor in the success of his studies. There is also an economic point of view here: at present, a great deal of our scarce resources must be spent on practical language teaching. Therefore, the better the students' initial language skills, the more resources can be directed towards attaining our other important objectives (it is well known that because of

this practical language component in addition to the theoretical matter to be learned, language studies are more demanding than other subjects in the humanities in terms of time). Moreover, as was pointed out above, we have at present no direct measures of many of the qualities that we should like to look for in our applicants. If we gave up the language skills as an important selection criterion, we would at the same time give up indirect evidence of some of the other desirable qualities mentioned above.

3. THE MATRICULATION EXAMINATION IN ENGLISH AS A SELECTION INSTRUMENT

The prevalent form of the matriculation examination, virtually unchanged since 1919, is a written translation from and into English. It falls far short of covering the modern goals of secondary school English teaching. Even those few modes of communication that are touched on, i.e. reading and writing, are assessed indirectly: through the ability to translate, which is, as is now generally (though not unanimously) agreed, a special high-level skill requiring, among other things, a good command of two languages but not unambiguously indicating proficiency in a foreign language. As far as the abilities to understand others and express oneself in spoken and written English are concerned, the translation test is an invalid selection instrument (for a discussion of its unsuitability for any other conceivable function and of its detrimental feedback on school instruction, see Hirvonen 1971a, 1-20; 1974, 34-66).

Technically, too, the translation test is unsuitable for selection purposes. Highly respected as it is by many teachers, there is no quantitative information available concerning its reliability. We just do not know whether the tests themselves or their marking are reliable.

The variance in the students' scores in the whole population is normally adequate, but the effective discrimination takes place at the lower end and the middle of the scale. Thus we get a distribution of marks with a very strong negative skew, indicating that the test is too easy to discriminate among the better candidates. This is clearly evident in the distributions of final marks during the past few years:

Percentages of final marks in the obligatory English tests of the Matriculation Examination, 1969-1973

i = improbatur
a = approbatur
b = lubenter approbatur
cl = cum laude approbatur
m = magna cum laude approbatur
l = laudatur
N = no. of candidates

year	i %	a %	b %	cl %	m %	l %	N
1973	5.18	13.64	16.34	20.90	21.08	22.86	16 595
1972	4.67	14.11	19.35	22.00	19.61	20.25	14 653
1971	5.12	12.32	17.74	22.31	20.49	22.03	12 844
1970	4.18	11.94	19.[illegible]	24.[illegible]3	20.30	19.70	11 101
1969	5.20	30.91		45.64		18.25	10 110

N.B. The intermediate marks b and m have been in use only from 1970 on.

(Being occupied with the autumn examinations, the Matriculation Board was unable to provide me with the 1974 figures in time for the deadline of this article. But I believe the above figures establish the pattern clearly enough.)

This mark Distribution, which does not make sense in terms of group-referenced or relative assessment in general (cf. Hirvonen 1974, 36-37), is of particularly low usefulness for the selection of students of English. With only the better candidates seeking admission in the first place and with the English departments being able to admit only the top 20 % of those (or less), it is clear that the results of the matriculation examination in English can give little or no help in discriminating between the applicants. In this group, a "laudatur" is nothing: everybody has one.

(I shall bypass here without further comment the unfortunate "experiment" with the so-called "new-form English test" in the matriculation examination in 1965-1973. For the reasons for this bypassing, see Hirvonen 1971a, 7-9 (footnote); 1974, 34-38.)

From 1974,on, an alternative examination is offered. Consisting of three parts - listening comprehension, reading comprehension, and an essay - it seems to be an improvement on the translation test. How well it will work remains to be seen. It will not, however, be seen until the new examination is made compulsory. At present, there is an utterly ridiculous system where the choice between the old translation examination and the new one is left to the individual candidate to make - in the

actual examination session (!). This system, which is to be continued for another couple of years on the pretext of "experimentation", effectively prevents meaningful experience from being gained about the working of the new examination and makes the results of the two examinations incommensurable, as the choice of examinations can be neither controlled nor statistically predicted. Probably the choice of the new examination indicates something about the quality of the instruction that the candidate has received, about his motivation for English, or both, but we have no way of taking this sort of vague information into account in our student selection.

I have tried to show above that the matriculation examination in English is an unsuitable selection instrument for English departments on account of both its lack of content validity and its poor technical character as a test. At least during the next few years, then, if English departments want accurate and representative information on their applicants' proficiency in the practical language skills, they must continue administering their own tests to assess it. Later on, if the matriculation examination in English is developed into a valid, representative, reliable and sufficiently discriminating instrument (and models and research results for such development are available even at present, cf. the bibliography), this issue may be reconsidered.

4. LANGUAGE TESTS AS SELECTION INSTRUMENTS: REQUIREMENTS AND LIMITING FACTORS

The requirements of efficient testing are discussed only very briefly here. More comprehensive treatment is to be found in any textbook on educational measurement or language testing.

1. Validity. The tests must measure what they are supposed to measure, and cover the objects of measurement in a representative way. This means that if the object of measurement is the candidates' ability to understand, read, speak and write English, the battery of tests must include representative sub-tests of listening comprehension, reading comprehension, oral expression and written expression. A testing procedure that does this is said to possess high content validity.

A test (or a battery of tests) that is used for selection should also have predictive validity, i.e. one should be able to demonstrate that those who do well in the test do well in their studies and/or their later

professional life, and vice versa. In the present situation, this is difficult to show because only the best applicants can be admitted and there is little variance in their test scores; in other words, we have no control group for a follow-up study. This is all the more reason to pay attention to the content validity of the entrance test.

2. Discriminatory power. The entrance test must discriminate among the applicants, i.e. one must be able on the basis of the test scores to arrange the applicants in an order of merit. When only about 20 % of the applicants can be admitted, the test must discriminate most effectively at this cutoff point, i.e. among the better applicants. This means that the test must be quite difficult.

3. Reliability. The test results must be dependable, i.e. variance due to random or systematic error must be minimized. This requirement will be appreciated by those who have sometimes had to make the painful decision of accepting a candidate with 59 score points and rejecting the next one who had 58 points. Quantitatively speaking, the reliability coefficient (e.g. as estimated by the Kuder-Richardson Formula 20) of a battery of tests used as a basis for selection of individuals should be upwards of .90 (see Mikkonen & Mikkonen 1971, 92).

4. Face validity and fairness. Besides being valid and reliable, the entrance test must look valid and appropriate to the candidates, be acceptable to them. Because even quite good candidates will have to be rejected in the selective entrance test, the administrative decisions must be justifiable to those who cannot be admitted.

Limitations on the selection procedures are imposed by the great numbers of applicants, the limited time available for testing, and the need to reach the decisions soon after the testing.

In 1974, the following number of candidates applied for admission to the English departments of the universities of Joensuu, Jyväskylä and Turku (the no. of those admitted is given in parentheses): Joensuu 120 (20), Jyväskylä 750 (80), Turku 642 (55). From the administrative point of view, numbers like this virtually rule out some testing techniques, such as interviews and even lengthy essays, and make a good case for the use of objective tests wherever applicable. For reasons of test security, each test must be administered in a single session, and there is a good case for having all the sessions on the same day, as it does not seem reasonable to expect the applicants to stop for several days for the entrance test. The month of June being reserved for sending in the applications the test cannot very well be administered until July. Both the departments and the applicants need the results as soon as possible: the departments in order to inform

the selected candidates, receive confirmations and cancellations, contact reserve candidates, receive more confirmations and cancellations, and so on until the quota is filled (hopefully by the beginning of term), and the candidates in order to make sensible plans for the academic year.

In conclusion, we need a representative, reliable sample of the candidates´ proficiency in the various language skills by means of tests that are quickly taken and quickly scored.

5. ASPECTS OF THE JOENSUU-JYVÄSKYLÄ-TURKU JOINT ENTRANCE TEST

For two years now, starting on the initiative of the English department at Turku, the English departments of the universities of Joensuu, Jyväskylä and Turku have selected their students by means of a joint entrance test that is compiled in co-operation and administered simultaneously in all the three universities.

As this joint entrance test is discussed in detail in Kohonen´s contribution, I shall limit myself here to a couple of comments on the advantages and disadvantages of the joint system in its present form.

Compiling the test in interdepartmental co-operation saves labour, and the combined expertise of three departments probably produces a test battery superior to what any single department could muster.

The joint procedure is economical from the applicant´s point of view as well. He need normally send only one application and take only one entrance test to be considered by three departments.
One intriguing aspect of the joint selection procedure is that, besides ensuring a uniform, high starting level as far as the students´ practical language skills are concerned, it may open the way in the future for joint achievement tests as well, e.g. "approbatur language skills", "cum laude language skills" and - why not? - in theoretical areas of study as well. Naturally, the eventual development of such joint examinations would have to be limited to those areas of study where close agreement on the objectives and the required levels of proficiency had been reached among the participating departments. Anyway, the advantages of such a system would be similar to those of the joint entrance test procedure: better examinations with less work and an assurance of a uniformly high standard of proficiency. If used with moderation and limited to the "common core" of university English studies, such joint examinations would leave ample scope for individual differences of

emphasis among the departments involved.

The battery of tests consists of sub-tests of vocabulary, structures, reading comprehension, listenin comprehension, and sound discrimination. The total administering time is about five hours. The testing technique used in all the sub-tests is multiple choice. With this technique, it has been possible to use hundreds of items in order to cover the objects of measurement representatively and to make up a reliable battery of tests. All the sub-tests can be scored and analysed by means of a computer direct from the candidates' answer sheets, and the scoring, analysis and publication of the results (= selection of students) is a matter of a couple of days.

As far as the coverage of the joint entrance examination is concerned, there are great gaps: oral and written expression are not tested. These are ruled out by the measuring techniques designed to cope with the practical problems of the number of applicants and the limited time and other resources. While the battery is, in my view, a tremendous improvement on the old selection procedure that limited itself to considering the applicants' translation ability (either in the matriculation examination or separately tested), it is nevertheless incomprehensive as yet. This incomprehensiveness can be partly (but only partly) justified by pointing out that candidates' performance in the receptive skills has been found to correlate fairly highly with that in the productive skills.

The language battery is not the only criterion employed in the selection. It accounts for 70 % of the maximum score. The other 30 % is constituted by the weighted sum of five other criteria:

1) School mark in English
2) Average of marks in the theoretical subjects in the school certificate
3) General mark in the matriculation examination
4) English mark in the matriculation examination
5) Mother tongue mark in the matriculation examination

In view of the theoretical aspects of university English studies and the present lack of dependable predictors of success in these (cf. pp. 6-7), I think it reasonable to give some credit for successful performance at school and in the matriculation examination. There is also some empirical justification for including the above criteria. In a study of the predictiveness of the Finnish matriculation examination, Lehtovaara (1966) investigated the extent to which the examination and its various part-examinations would predict success in studies at the University of Helsinki, measured by a combination of four criterion variables:

a) no. and quality of degrees in the master´s certificate (FK)
b) mark given for the 'pro gradu' thesis
c) the time needed to take one´s degree
d) qualitative mark received in the main subject

The students´ results in the matriculation examination correlated with success in English Philology as main subject in the following way (only the correlations for women are given here, since there were so few men studying English Philology that the sample did not seem very representative; source: Lehtovaara 1966, 134-160):

1.	Total points of merit (max. 18)	$.48^{+++}$
2.	Total points of merit, obligatory tests only (max. 12)	$.48^{+++}$
3.	General mark (a = 1 c1 = 2 1 = 3)	$.41^{+++}$
4.	Mother tongue (-"-)	$.40^{+++}$
5.	Other national language (-"-)	$.36^{+++}$
6.	Foreign language (almost exclusively English in this case) (-"-)	$.24^{+}$
7.	General knowledge ('reaali') (-"-)	$.43^{+++}$

Note: +++ = statistically significant at the risk level $p < .1$ %
+ = statistically significant at the risk level $p < 5$

According to this study, the value of the various tests in the matriculation examination is not insignificant in predicting success in English studies (these results are relevant, of course, only to the extent that the objectives, areas of study, and criteria of success are the same as those in English Philology at the University of Helsinki in the late 1950´s and early 1960´s). The best predictor seems to be the total points of merit either in all subjects or in obligatory subjects of the matriculation examination. The least significant predictor is the foreign language test (= translation)(!) (One explanation for this may be that this test had been an important criterion in the selection of students, so that there was little variance left among those selected.)

It seems to me that 30 % is quite enough for the school record and the matriculation examination results, partly because there are doubts regarding their validity, and especially because their reliability is unknown.

6. SUGGESTIONS FOR FUTURE DEVELOPMENT OF STUDENT SELECTION

It would seem profitable to all parties concerned to extend the co-operation in student selection to include all the English departments that have similar objectives and similar selection criteria. By means of computer technology, the joint procedure can easily be made flexible enough to permit a participating department to employ additional selection criteria of their own, if they so desire. Besides, each department can have a say in developing the joint procedure itself into a more versatile and efficient one. At present, the procedures used are certainly not ideal, and the three departments who started the joint effort would welcome new "members" with fresh ideas.

In its present form, this joint entrance test procedure seems to select students with impressive school records and matriculation examination results and a good structural grasp of English. Many of them, however, have such inhibitions, especially regarding the oral use of English, as to render their communicative competence greatly inferior to their linguistic competence. Indeed, several students have initial difficulty in expressing themselves in Finnish, let alone in a foreign language. Now, with the indisputable need for a shift emphasis in university instruction from the traditional mass-lecture-type learning situations towards group-centered, discovery-oriented activities with a collective responsibility for learning tasks, we need students who are able and willing to contribute to the learning situations, who have something to say and the confidence to say it. Clearly, the development of self-confidence and of the ability to communicate and to function in a group should be among the outcomes of university training, but it does seem reasonable to expect that we had something to go on initially. While it is neither possible nor desirable to fill our departments with raving extroverts, I cannot see why we shouldn't try to get students who have the necessary self-confidence to speak - preferably in English. This is one thing our present selection procedures are not doing.

The above considerations speak strongly in favour of inclusion of an oral test, an interview perhaps, in the entrance test battery. I appreciate the administrative difficulties this would lead to if all the applicants were interviewed, but I suggest that we investigate the feasibility of a two-stage system where the applicants are short-listed, as it were, by means of a test battery of the present type and the final selection is made on the basis of an interview.

As was suggested on pages 6-7, we lack a valid and dependable measure of the applicants' "verbal intelligence" and of their problem-solving ability. Yet these abilities should, in my view, be important criteria in the selection. In the near future, we might try to work out tests of verbal reasoning. They could take the form of a limited corpus of material in an invented language followed by linguistic problems to be solved on the basis of the data. At the experimental stage, these sub-tests could be given very little (or even no) weight in the selection until their validity and reliability were established.

If the experiments with tests and measures of important prerequisites other than practical language skills are successful, the question may arise whether the main discrimination among the applicants should still be done on the basis of their language skills. Instead, one could think of a system where the necessary linguistic competence would be defined beforehand in terms of demonstrable performance criteria. The applicant would simply pass or fail this criterion (i.e. our group-referenced tests would be replaced by criterion-referenced ones), and the selection among those passing would take place on the basis of their linguistic problem-solving ability and their communicative competence (not necessarily language-specific).

Finally, a word about the requirement of equal chance. It has been argued that when language skills, particularly oral language skills, are emphasized in the selection, injustice is done to those applicants who, while possibly highly talented, have received inferior instruction at school (remote location, incompetent teachers, insufficient materials and technical aids etc.). On the other hand, those applicants whose parents are wealthy enough to send them abroad during summer holidays are said to have unfair advantage in the entrance test. I think it is true that the chances of admittance through the entrance test are unequal. But this is just another way of saying that those who are highly proficient in English stand a better chance than those who are less so. Now where exactly does the injustice lie? I think injustice is done to pupils wherever they are given inferior instruction at school. When this is the case, money will be spent to buy more or better "instruction" - private lessons, tape-recorded materials, foreign travel .... Not everyone can afford this, of course. It is not in the power of university language departments to re-distribute the wealth in our society, but if the general standard of foreign language teaching at our schools could be raised and unified, the effects of social and economical inequality would diminish in proportion. This is something where the universities can contribute by training uniformly

better teachers and researchers. To succeed in this, we need to get the best "raw material" available. In this situation, I cannot see whom we might be thought to help if we used lopsided, inefficient entrance tests so as to be able to ignore the fact that the quality of instruction in our schools varies and to prevent ourselves from selecting the best qualified applicants. Technocratic as this may sound, it seems to be the only solution that makes sense so long as our scarce resources force us to be highly selective.

SELECT BIBLIOGRAPHY

Black, Colin & Viljo Kohonen (1972): Englantilaisen filologian oppilasvalinnat 1971, Turun Yliopiston Englantilaisen filologian laitoksen julkaisuja n:o 2, Turku 1972

Davies, Alan (ed.)(1968): Language Testing Symposium, London 1968: Oxford University Press

Hirvonen, Pekka (1971-1974): Englannin kielen taidon mittaaminen lukion päättyessä:
(1971a): 1. Lähtökohdat, Suomen sovelletun kielitieteen yhdistyksen (AFinLA) julkaisuja n:o 1, Turku 1971
(1971b): 2. Kuuntelukoe, AFinLA:n julk. n:o 4, Turku 1971
(1973a): 3. Lukukoe ja Kirjoituskoe, AFinLA:n julk. n:o 7, Turku 1973
(1973b): 4. Puhumiskoe, AFinLA:n julk. n:o 8, Turku 1973
(1974) : 5. Kielitaitokoe, AFinLA:n julk. n:o 9, Turku 1974

Hirvonen, Pekka (1972): "Miten lukion kielten loppututkintoa olisi kehitettävä", Tätä mieltä kielten opetuksesta (toim. Pekka Hirvonen), AFinLA:n julk. n:o 5, Turku 1972, ss. 54-66

Kolmannen asteen kieltenopetuksen nykytila, Lingvan tutkinnonuudistusmateriaalia 1/73 (moniste), Lingva r.y. 1973, Helsinki

Komiteanmietintö 1972: A 17: Filosofisten ja yhteiskuntatieteellisten tutkintojen toimikunnan mietintö

Komiteanmietintö 1974: 101: Vuoden 1973 opettajankoulutustoimikunnan välimietintö

Lehtovaara, Arvo (1966): The Predictiveness of the Finnish Matriculation Examination, Suomalaisen tiedeakatemian toimituksia, sarja B, nide 140, Helsinki 1966

Mikkonen, Valde & Juhani Mikkonen (1971): Opsam: Opintosaavutusten Mittaus, Helsinki 1971: Tammi

Mikkonen, Valde (1972): Kieli-instituuttien oppilasvalinnat 1971, AFinLA:n julkaisuja n:o 6, Turku 1972

"Opetusministeriön kannanotto 13.3.1974 korkeakoulujen tutkinnonuudistuksen jatkosuunnittelusta", Korkeakoulu- ja tiedeosaston tiedotuslehti 3/1974, Opetusministeriö 1974, Helsinki

Otter, H.S. (1968): A Functional Language Examination, London 1968: Oxford University Press

Ylioppilastutkintoasetus ja yleisohjeet ym., Ylioppilastutkintolautakunta 1973, Helsinki

Viljo Kohonen:

JOINT ENTRANCE TEST JOENSUU-JYVÄSKYLÄ-TURKU, 1974: SELECTION PROCEDURES AND TEST ANALYSIS

This paper covers the following aspects of the project of a joint entrance test for the English departments of the universities of Joensuu, Jyväskylä and Turku, carried out in summer 1974: planning, administration, data processing on the computer, selection of students, and a statistical analysis of the test battery, with some examples of the item analysis. The co-operation between these universities began in summer 1973, when a basically similar, jointly prepared entrance test was administered simultaneously by the three English departments; for the background of the project, see previous publications by Black and Kohonen (1972) and Kohonen (1973).

1. PLANNING AND ADMINISTRATION

1.1. Planning. The team in charge of the planning and administration of the entrance test consisted of Pekka Hirvonen and Jorma Tommola from Joensuu; Glyn Hughes, Peter Jarrett, Raija Markkanen and Anthony May from Jyväskylä; and Keith Battarbee, Risto Hiltunen and Viljo Kohonen from Turku. The planning was started in November 1973, and a major part of the work was carried out by letter and by telephone, the first joint meeting being only in April 1974, when the skills to be tested, and, accordingly, the types of tests to be used, were agreed upon, and the amount of work divided between the project group. The proposed entrance test was submitted to the departments in late May, and in its final form the examination consisted of 12 tests. The numbers of items in each test, with number of choices and maximum scores, and the weight coefficients were as follows:

	items	max	no. of choises	weight coefficient
1. overall grade from the matriculation examination	1	9	-	3
2. grade for English from the matriculation examination	1	9	-	1
3. grade for native language from the matriculation examination	1	9	-	3
4. average of academic subjects in the school-leaving certificate	1	9	-	4
5. Teacher's grade for English in the school-leaving certificate	1	9	-	1
6. vocabulary and grammar I	71	71	4	7
7. vocabulary and grammar II	76	76	3	6

8. translation analysis	50	50	4	5
9. reading comprehension	26	26	4/5	5
10. auditory discrimination	50	50	3	1
11. listening comprehension	26	26	3	3
12. lecturette in English	10	10	3	1
total of items and max. score	314	354		

Tests nos. 6 and 7, and part of 9, were prepared in Turku; nos. 8, part of 9 and 11 and 12 in Jyväskylä; Joensuu was in charge of no. 10. All the items were checked and commented upon by the other working groups before the final version was arrived at; in many cases, this cross-checking meant a considerable improvement to, or a complete revision or deletion of, a suggested item. According to the agreed weight coefficients, school performance accounted for 30 per cent of the total score and the test battery for 70 per cent. In determining these coefficients, account was taken of the length of the test as well as whether it measured a sub-skill or a total skill. Thus, for example, test no. 10 (auditory discrimination) was given a coefficient of only 1 in spite of its length, because it measured a sub-skill. These coefficients also resulted in the listening tests (nos. 10-12) representing altogether about one fifth (18 per cent) of the test battery (tests 6-12).

All the papers were offset-duplicated in Turku (a print of 1,000 copies, total of 30,000 sheets of paper) and stapled automatically in booklets of 30 pages, one for each candidate taking the test.

1.2. Information. Advance information for candidates was included in each university's printed study guides, published in spring 1974. These guides explained the nature of the examination and the process of application. With the exception of Jyväskylä, which required all its candidates' applications, candidates were requested to send their applications only to the university of their 1st preference, stating in their applications whether or not they wished their scores to be taken into consideration in the other two universities. Notwithstanding this, they could take the entrance test in any of the three places. As it was essential for each department to know how many candidates they were going to have, the applicants were obliged to return the registration form which they had received together with the advance information (or the application form) to that university where they intended to take the examination (see Appendix 1 for the advance information and registration form for Turku, in Finnish.) The aim of this method was to reduce candidates' travelling expenses by permitting them to take the test at the university which was nearest their home. This process worked as envisaged: in each place, there were several candidates applying to the other two universities as their 1st preferences.

1.3. Preparations for computer scoring. The results were scored on a Univac 1108 computer using a library programme called OPSAM (for the manual, see Mikkonen and Mikkonen 1971), particularly designed for an objective scoring of tests. All the items in tests 6-12 were therefore multiple-choice items, directly suited for computer scoring. The candidates' school performance, however, had to be converted into one digit data, as specified by the programme. This was done by the planning team after the examination by using the following conversion tables:

Matriculation examination:

grade	before 1969	after 1970
l	9 points	9 points
mcl	-	7
cl	6	5
b	-	3
a	2	1
i	0	0

Average of academic subjects:

9.5 -	9
9.2 - 9.4	8
8.9 - 9.1	7
8.6 - 8.8	6
8.3 - 8.5	5
8.0 - 8.2	4
7.7 - 7.9	3
7.4 - 7.6	2
7.1 - 7.3	1
- 7.0	0

The teacher's grade for English was converted into one-digit data by subtracting 1 from every score (range 3 -9).

To get rid of the cumbersome card-punching, candidates were requested to mark their answers onto the answer sheets of an optical reader, available in Jyväskylä; the cards needed for OPSAM were punched by the optical reader. A small conversion programme was, however, necessary before the data, punched by the optical reader, were in accordance with the OPSAM specifications. This programme, as well as the computer run-through, were taken care of by Arja Manninen, while Raili Rossi was in charge of the optical reader.

1.4. Administration of the test, on July 8th. In order to standardize the examination conditions, all the test papers were edited and duplicated by the same working group (Turku), the examination commenced at precisely the same time and lasted the same time (written tests, 10 - 14; listening tests, 15 - 15.50). The duplicated instruction sheet (see Appendix 2, in Finnish) was given to each candidate and explained at the beginning of the examination. At the same time, the following practice example was done together, with instructions for marking the answer on the optical reader answer sheet:

Practice example:

1. He could not make his voice _____.
   a) being heard b) to be heard c) heard d) hear

Answer as follows:

1. A B C D E

Draw, with a pencil, a thick line below the intended alternative, between the broken line.

The listening tests were given in lecture halls so equipped with loudspeakers that listening conditions were identical in all parts of the hall. Here, too, the instructions were identical in every place, as everything has been recorded on tape by the Jyväskylä group, and copies had been made for Joensuu and Turku. Thus, the examination conditions can be said to have been comparable.

To find out student opinion about the tests and the selection procedures, a questionnaire (prepared by Väinö Roininen) was given after the examination to a sample of every third candidate taking the test. The results of the questionnaire are discussed in the paper by Roininen and Kosonen.

1.5. Scoring on the computer. All the answer sheets (5 per candidate, altogether 4,215 sheets) were taken to Jyväskylä, where the school performance data (sub-tests 1-5) were transferred by the planning team onto the answer sheets by using the above mentioned conversion tables. The punching of the cards on the optical reader was completed on the evening of July 9th, when it was discovered that some 30-40 per cent of the cards were not punched properly: the optical reader had not been able to read all the answers, since many candidates had used too hard and sharp pencils, not drawing thick lines between the broken line, as instructed; the optical reader failed to read and punch all such answers. This meant two days' work for the planning team and other members of the staff at Jyväskylä, because all such answer sheets had to be "corrected" by hand, i.e. the candidates' markings had to be made clearer with a soft, blunt pencil. After this, the faulty cards were re-punched, and the results were available on the 13th of July, belated by one day from the deadline announced in advance.

2. RESULTS

2.1. Statistical analysis of the tests. The main results of the statistical analysis of the test battery (completed by 843 candidates) are summarized in Table 1: original no. of items in each test; no. of acceptable items (i.e. those that were shown to discriminate properly); range of scores after rejection of poor items, with means and standard deviations of the scores; and reliability values for the tests and the test battery (for the mathematical formulae, see the manual of Mikkonen and Mikkonen 1971).

Table 1. Main results of the analysis of the tests.

test	no. of items	accept. items	range of scores	$\bar{x}$	S.D.	reliability
6	71	70	10-65	34.6	9.8	.86
7	76	67	12-63	45.3	6.2	.70
8	50	49	12-45	30.6	5.7	.74
9	26	26	2-26	16.3	4.8	.79
10	50	50	17-50	40.1	7.4	.89
11	26	26	5-26	18.9	4.3	.78
12	10	10	1-10	5.7	2.1	.54
whole test	314	303	124-303	222.7	33.4	.91[1]

Another measure of the difficulty level of the tests is provided by the average solution percentages on each test. These, with the range of the percentages and standard deviations, were as follows.

test	range	$\bar{X}$	S.D.
6	5.1%-96.9%	48.8%	22.2
7	12.1 -98.7	64.2	22.4
8	4.3 -94.3	62.1	24.0
9	30.4 -80.3	62.7	13.8
10	43.4 -97.1	80.2	13.7
11	46.1 -91.8	72.7	13.3
12	32.5 -77.5	56.8	14.4

It can be said that the test battery functioned well: only 11 items were discarded because of poor discriminatory power, and the wide range of scores and solution percentages shows that there were both easy and difficult items; tests 6, 7 and 8 have some very difficult items, which mark off the élite as well. This is necessary, considering that only some 18 per cent of the candidates could be admitted. The reliability values for individual tests (alpha coefficients, indicating internal consistency of the test) are satisfactory, and the reliability index for the test battery (.91) must be considered as being

---

[1] These totals also include sub-tests 1-5 (school performance). As these were scored using a scale 0 - 9 (see above), the maximum score of the accepted items was 343; the highest score achieved was 303 (= 88.3 %). Numbers of tests refer to: 6 = vocabulary and grammar I, 7 = vocabulary and grammar II, 8 = translation analysis, 9 = reading comprehension, 10 = auditory discrimination, 11 = listening comprehension and 12 = lecturette in English.

good[1]; .90 is generally held to be the minimum requirement in making important decisions.

2.2. Item analysis. In the OPSAM computer programme, the quality of an item is examined by a correlation procedure, where the correlation between the correct answer and the subtest should be positive, and those between the other choices (distractors) and the sub-test, negative, if the item is to function consistently, i.e. superior candidates solving it correctly more frequently on average than the poorer candidates. Consequently, any item failing to show a positive correlation (cut-off point normally $\geq .00$) is rejected by the computer, and the scores on such items are subtracted from every candidate's final scores. This is justifiable, since in such a system nobody is "penalised" for the test designer's failure to make unambiguous, well-discriminating items, and the important decisions are based only on the information yielded by items that function consistently in the population and have a decent discriminatory power. In multiple-choice items, the averages of the item correlation generally range between .15 and .30; they may be higher in standardized tests, where poor items have already been "pruned" from the test. In the present population ($N = 843$), $r = .09$ is signifant at the 95 % level, and $r = .12$ at the 99 % level (Mikkonen and Mikkonen 1971, 136).

These principles can be illustrated best by some examples of the item analysis.

Vocabulary and grammar I, item 41:
I'm afraid we must _____ our bookings for the play today.
a) remove b) reverse c) unbook d) cancel

Result: d) r= .46, solution percentage 50.3% [2]
a) r= -.38, N= 204 (=24.1%);
b) r= -.21, N= 122 (=14.4%)
c) r= -.17, N= 92 (=10.9%)

The high positive item correlation (.46) shows a very good discriminatory power, while the negative correlations of the distractors indicate that they had been chosen by poorer candidates, and the answers were distributed over all of the alternatives.

---

[1] The reliability index of the test battery can be interpreted as the probability of the result being the same if the examination were to be repeated. The lower reliability value of test 12 (lecturette) is largely due to the small number of items (only 10), and/that the items had no more than 3 alternatives; in such a case, reliability is affected by chance variation more than in a longer test.

[2] Symbols: r = correlation; N = number of candidates who chose the particular alternative, and percentage of these out of the whole population; + (asterisk) = the item was discarded by the computer because of failure to discriminate properly. The correct alternative is mentioned first, with percentage of correct answers.

Item 66:
The government resigned, but continued in office as a(n) _____ government.
a) executive b) caretaker c) substitute d) vice

Result: b) +r= -.06, solution percentage 8.3%
a) r = .02, N = 263 (= 31.1%);
c) r = -.05, N = 305 (= 36.1%);
d) r = -.10, N = 202 (= 23.9%)

This item failed to discriminate, since it was too difficult even for the better candidates, two of the distractors being more popular than the correct answer (b). The distribution of the choices over the alternatives was thus at a random level, and the item was discarded by the computer, as indicated by the asterisk before the correct choice.

Vocabulary and grammar II, items 95 and 96:
Born on August 4, 1900, the Queen Mother was the youngest daughter of the earl of Strathmore, a Scottish peer. She was the first bride

(95) a) having no
b) not of / royal birth
c) exempt from

(96) a) having been chosen
b) to become chosen / in
c) to have been chosen

many years by the son of an English King.....

Result: (item 95) b) r = .28, solution percentage 45.7%
a) r = -.38, N = 414 (= 49.1%);
c) r = -.06, N = 17 (= 2%)
(item 96) c) +r= -.05, solving percentage 27.4%
a) r = -.02, N = 271 (= 32.1%)
b) r = -.15, N = 339 (= 40.2%)

No. 95 is a good item, even though alternative (c) had attracted only 17 candidates, and should perhaps be revised to make it more attractive. Item 96 was discarded by the computer because of inconsistent discrimination. Alternative (b) seems all right, but many of the better candidates were distributed between alternatives (a) and (c). It may be that the item is ambiguous because the borderline between the gerund and the infinitive is vacillating in spoken language.

Translation analysis, item 189:
Hän odottaa (=wait), että mies nyt sanoisi jotain.
a) for the man to say b) that the man would say
c) the man to say d) for the man should say

Result: a) r = .32, solution percentage 51.7%
b) r = -.32, N = 270 (= 32%)
c) r = -.08, N = 72 (=8.5%)
d) r = -.27, N = 48 (=5.6%)

The item is a good one. Alternative (b) attracted about a third of all choices, and the explanation is clearly

contrastive: the poorer candidates "translated' the että clause word for word, failing to recognize the English for+accusative+infinitive structure.

Item 190:
Mies havahtuu, kuin olisi miettinyt jotain.
a) he thought b) he would have thought
c) he was thinking d) having thought

Result: c) r = .17, solution percentage 10.8%
a) r = .01, N = 25 (= 2.9%)
b) r = -.28, N = 378 (=44.8%)
d) r = -.00, N = 343 (=40.8%)

The item is very difficult and discriminates fairly well. Alternative (b) is a good distractor, based on a contrastive word-for-word translation, but (d) is dubious: it was chosen by 40.8 per cent of the candidates, and the zero correlation indicates that these included both good and poor students, and the same also applies to (a). The item is in need of revision. It is also indicative of the difficulty in "cooking up" plausible distractors: it is frequently difficult to invent more than one or two good distractors, and in such a case there is not a real choice between 4 alternatives, but perhaps only two. This is particularly true about items intended to measure tense or aspect, as here.

Reading comprehension, item 203:
Man is inconsistent in his likes and dislikes. What he loves at 20 he hates at 40, but this inconsistency speaks well of his ability to profit from experience. If he never changed his attitudes, he would be marked as.... .
a) a wise man b) preferring simple things
c) unable to learn d) extremely inconsistent

Result: c) r = .38, solution percentage 72.3%
a) r = -.14, N = 36 (= 4.2%)
b) r = -.17, N = 49 (= 5.8%)
d) r = -.37, N = 129 (=15.3%)

The item has a good discriminatory power, even though it is a fairly easy one. As the listening tests are discussed in detail in the paper by Hughes and May, they are not analyzed here.

2.3. School performance. Performance at school (sub-tests 1-5) was converted to the scale 0-9 (3-9 for the teacher's grade in English), and the analysis gave the following distributions for these:

test	range	$\bar{X}$	S.D.	rel.	$\bar{X}$ corresponds to:
1	1 - 9	6.8	1.7	-	c1 - mc1
2	1 - 9	7.4	1.9	-	mc1 - 1
3	1 - 9	6.0	2.2	-	c1
4	1 - 9	3.5	2.3	-	8.0 - 8.2
5	5 - 9	7.5	1.0	-	8 - 9

It will be seen that the averages are quite high, particularly in tests 2 and 5 (English in the matriculation examination and the teacher's grade for English), i.e. the discriminatory power of the school grades is fairly poor; in English, at least, there is a problem of test ceiling: the grades do not discriminate the better candidates, which makes them statistically of little importance. This is not to say, however, that school grades have no discriminatory value in general: in an ordinary classroom there is a fair amount of discrimination, no doubt, and it is only when the better candidates of various classroom populations come together that the language grades particularly lose their discriminatory power. There are no data on the reliability of the school grades, since the computer programme is not able to calculate any index on the basis of one item only.

2.4. Intercorrelations of the tests. The intercorrelation matrix gives a possibility for examining connections between the tests.

	1	2	3	4	5	6	7	8	9	10	11	12
2	.59											
3	.57	.24										
4	.65	.34	.41									
5	.57	.56	.32	.59								
6	.38	.38	.19	.30	.48							
7	.35	.29	.23	.23	.40	.64						
8	.36	.33	.22	.28	.44	.64	.55					
9	.37	.32	.29	.31	.40	.49	.55	.47				
10	.10	.11	.06	.04	.13	.26	.32	.28	.36			
11	.17	.14	.14	.06	.19	.37	.44	.39	.48	.76		
12	.24	.21	.19	.16	.27	.38	.44	.33	.49	.50	.56	
(13)	.53	.47	.37	.41	.57	.81	.78	.74	.74	.61	.69	.62

All the correlations are positive, and the overall grade in the matriculation examination has high correlations to all the school grades. Among the tests, there are high correlations between the vocabulary tests and the translation analysis (nos. 6, 7 and 8), indicating that there is some overlapping in their measuring objects. Listening tests also have quite high intercorrelations, forming a group of their own ("listening factor"). Three factors seem to emerge from these correlations: school performance, written tests and listening tests.

2.5. Selection of students. Of the 843 candidates who completed the examination, 84 gave Joensuu, 376 Jyväskylä, and 383 Turku as their first preference. The computer was programmed to produce the results, based on the absolute final scores, on a relative scale of 0 - 80, on which the cut-off points turned out to be 62 for Turku, 54 for Jyväskylä and 53 for Joensuu. The numbers of students admitted

to the three universities were, according to the quotas decided by the faculties, Turku 59 (= 15,4 % of those who gave Turku as their first preference), Jyväskylä 86 (= 22.9 %) and Joensuu 18 (= 21.4 %), total 163 (= 19.3 % of those who completed the examination).[1] The checking of the candidate second and third preferences was done by hand, and was relatively straightforward: for instance, if a candidate failing to get admission to Turku was shown to have a score that placed him above the borderline for Jyväskylä and Joensuu, it was established whether he had indicated his willingness to be accepted by either of these universities, and if so, he replaced some other candidate at the lower end of the scale.

The Turku intake was examined more closely in terms of the school background. The averages of the school grades of these 59 students were shown to be considerably higher than those of the whole population, as one could expect:

Turku intake (N = 59)			Whole population (N = 843)
1	8.8	= mcl - 1	6.8
2	8.2	= mcl - 1	7.4
3	7.9	= mcl - 1	6.0
4	5.9	= c - mcl	3.5
5	8.5	= 9 - 10	7.5

The differences were not tested statistically, since the averages for the whole population also include the candidates that were admitted to the three universities, but the figures are clear enough to show that those who were admitted to Turku, for instance, had very high school grades, which makes the problem of test ceiling even more acute with the best fifth of the candidates, i.e. those among whom the quotas are filled.

Another problem is posed by candidates who have done the short course in English at secondary school (3 years, or the "half-long" course, 6 years, in secondary schools for girls). In the Turku population (N = 383), there were 46 such candidates, and 4 of them had a score above the cut-off point of 62, i.e. 8.7 per cent. It seems that the test is probably too difficult for these candidates, since the four accepted candidates had the "half-long" school background in English. On the other hand, the standard of these candidates was not very high (the above averages were, respectively, 6.7; 7.2; 5.8; 3.4 and 7.6).

---

1 These figures also include the reserves. The quotas were Joensuu 20 (two places had been already filled in May), Jyväskylä 80 and Turku 55.

3. DISCUSSION

If we review the test battery in the light of the essential principles of any testing - validity, reliability, discriminatory power, economy and fairness - it may be said that these requirements are met quite well. The test battery covers a variety of measuring objects (school performance, grammar, vocabulary, ability to analyze a Finnish sentence context, reading comprehension, listening tests) and has thus a decent content validity, even though it is still an open question whether tests of the production of written and spoken language should be included, and to what extent the tests have predictive validity (cf. Hirvonen's paper).

The reliability of the tests was shown to be good (whole battery .91), and the same applies to the discriminatory power of the tests; here there was a clear difference between the school performance (sub-tests 1-5) and the entrance tests (6-12): as far as the best fifth of the candidates was concerned, the school grades had very little discriminatory power, while the entrance tests were difficult enough to discriminate these candidates also (the highest absolute score, which included the school grades as well, was 303 out of the theoretical maximum of 343, i.e. 88.3 %). The considerations of economy were met by the use of computer technology in the scoring of the tests; it is worth noting that, in spite of some difficulties with the optical reader, the results were available to the candidates, with full information about their performance in each sub-test, only five days after the examination.

All these considerations may be seen as part of the problem of fairness. In entrance tests, the testers will make decisions that affect the whole career of some 20 000 young people every year, as far as university education is concerned in Finland, and the requirements of fairness are quite justified, considering that only some 50 per cent of these can be accepted for the different universities. In the English departments, the problem is even more acute, because only some 20 per cent of the candidates can be admitted. This means that, at the cut-off points at least, we are rejecting candidates that are in practice as good as many of those who were admitted, and in such cases we should be able to justify our decisions in as concrete terms as possible: the candidates should have full advance information about the selection procedures and principles, the tests should meet the normal technical requirements, and the results should be public.

As far as administration is concerned, it will be essential to emphasize to the candidates the importance of marking the answers clearly on the optical reader answer sheets. This should already be done in the advance information

to be given in the printed study guides, where every candidate taking the test should be requested to have a couple of soft pencils for marking the answers, and the technique must be rehearsed in the actual examination.

In its present form, the test battery is probably too long, i.e. the same information about the performance of the candidates could be obtained by means of a shorter examination, particularly as there seems to be overlapping between some of the tests (school performance; written tests; listening tests). The implication is that some of the present tests could be removed from the battery, and new ones introduced instead without making the examination longer. This will open an opportunity to experiment with totally new types of tests, in a population that is highly motivated and statistically very representative. The possibilities for developing the system are certainly far from being exhausted yet.

FOR FURTHER READING

Black, Colin and Viljo Kohonen, Englantilaisen filologian oppilasvalinnat 1971. Turun yliopiston Englantilaisen filologian laitoksen julkaisuja n:o 2, Turku 1972.

Corder, S. Pit, Introducing Applied Linguistics. Penguin Education 1973.

Harris, David, Testing English as a Second Language. New York 1969.

Ingram, Elisabeth, "Item Analysis", in: Davies, Alan (ed.), Language Testing Symposium, Oxford 1968, 192-203.

Ingram, Elisabeth, "Language Testing", in: Allen, J.P.B. and S. Pit Corder (eds.), Techniques in Applied Linguistics, The Edinburgh Course in Applied Linguistics volume 3, London 1974.

Kohonen, Viljo, "The Departmental Entrance - or a Joint Entrance Examination?", Kielikeskusuutisia 3/1973 (Jyväskylä), 1-11.

Lado, Robert, Language Testing. London 1961.

Magnusson, David, Testteori. 3rd edition, Stockholm 1969.

Mikkonen, Valde and Juhani Mikkonen, OPSAM, Opintosaavutusten mittaus. Helsinki 1971.

Pilliner, A., "Subjective and Objective Testing", in: Davies, Alan (ed.), Language Testing Symposium, Oxford 1968, 19-35.

Valette, Rebecca, Modern Language Testing: A Handbook. New York 1967.

Wright, David, The MCT: What Are We Testing? Project MUP, Rapport 7, Engelska Institutionen, Göteborgs universitet, 1969.

Peter Jarrett:

OBSERVATIONS ON THE 1974 TRANSLATION ANALYSIS TEST

I shall examine first the reasons for using this method of testing, then indicate the areas covered by the test, and finally look at students' performance and try to make some tentative conclusions.

1. The advantages of using a multi-choice testing method, apart from the obvious but secondary one of being able to put it through a computer, are, I think, the following. Firstly, one can pinpoint a specific problem, and eliminate the need for the often long contexts of the traditional translation text, which are time-consuming for both student and examiner and usually give very little information about the student's ability in any case. This, at least was our experience in early entrance exams in Jyväskylä, and prompted a gradual change in the direction of the present form. Secondly, there is the advantage of knowing the point of each item - they have been chosen on the basis of errors made by students during their initial stages of study. Thus the items are realistic, not based on supposition alone. Thirdly, a large number of problem fields can be covered in a relatively short test - 50 items in about 40 minutes, say. To cover the same problems with a traditional test form would involve a text or texts of enormous length, which it would be impossible to administer in the time available. Moreover, although one might learn a great deal about factors like the students' command of vocabulary in that way, it could more economically and logically be done by means of a specific vocabulary test.

The traditional type of translation test poses problems of selection which are in fact frequently ignored. (They need not be of course, but the multi-choice type of test certainly concentrates the mind of the examiner on these aspects automatically). If one asks how they are usually chosen in practice, the conclusion is that something called vaguely the 'difficulty' of the passage, or its 'suitability' is referred to, and in fact this usually means an impressionistic assessment of the difficulty or otherwise of the vocabulary in the text in question. The other common criterion is that the passage forms a whole as regards its subject, i.e. makes 'sense' as a separate unit. (There is also sometimes an attempt to choose an interesting subject, whatever that might be.)

These criteria, however, are largely negative, and come into operation only when one has already got to the stage of looking at texts, whereas what is needed is a

positive conception of what should be tested before one begins the actual process of selection or compilation. [1]

2. One must have a reasonably clear idea of what one is looking for and testing. What structures should be tested - futures, continuous tenses, articles, word order, etc., and how they affect the meaning of the sentence? Interference? If so, where, and what kind? Vocabulary? If so, of what type and register?

It seems to the writer to make sense in this test to concentrate on those aspects, as evinced by actual errors, which seem to be caused by interference from the native language, and especially structures. Other difficulties of structure and vocabulary are better dealt with elsewhere, and are in fact so covered in the present battery of tests.

Accordingly, this test included items in the following areas:
expressions of the future, various Finnish conditional structures, simple and continuous tense distinctions, perfect tenses, passives, word order (including questions and indirect speech), and the uses of sellainen, kukaan... ei, -na, -nä, etc. Not much emphasis is placed on vocabulary. The test mostly supplies the vocabulary for the student since a) it is tested in tests 6 and 7, b) subsequent analysis is made easier by eliminating complicating factors, c) it was felt to be fairer to students with less exposure to English, e.g. those with 'lyhyt englanti' (i,e. 3 years of English in the secondary school) or non-matriculated students, etc., and d) a command of basic structures should be expected, whereas the significance of any individual vocabulary item is obviously limited unless it is a particularly frequent error.

---

1. The matriculation exam (old type, translation option, F-E) also largely follows this pattern, as far as one can see; it is made up by the examiners to fit the English passage, of course. In general, it is not possible to see any precise motivation for the selection of difficulties occurring in the texts, except to make sure that the vocabulary is fairly familiar, which, as I have said, is a negative selection factor. There are inconsistencies between one text and another - on some occasions there may be a few structural difficulties (presumably still heavily penalised); on others, none at all - one can sometimes simply translate word by word, with no change in word order or structure. One text this year involved virtually little more than this. This is not a test of translation, but of vocabulary, and is ineffective even as that.

3. If one examines the performance of students in the main areas covered it emerges that on the basis of this test at least students performed relatively poorly in the items involving various kinds of passive constructions (selitetään, ajettiin, etc.), particularly in determining the subject of the verb. Of those areas reasonably well covered in the test the solution percentages were lowest here (even as low as 4 % in one case, but there were other items with 25 %, 33 % and 38 %, and only two fairly high solution percentages). Performance in items concerning the equivalents for the various Finnish conditional types was uneven, though the more common types were done well - a predictable result, perhaps. Futures were well done, but perfect tense items showed some weaknesses (present perfect-simple past distinctions where Finnish and English usages differ, for example). In regard to word order the position of words like 'also', 'already' and 'hardly ever' was very poorly understood indeed.

On the whole, and bearing in mind the fact that such items did not occur in all aspects of structure tested, the distractors which proved most 'effective' were those where the Finnish construction was followed most closely, i.e. literal translations. This was especially the case with verb tenses and word order. For example it would seem, on the basis of this test and other experience, that wherever the word myös is placed in a Finnish sentence, that position is the one most likely to be chosen by students in translating that sentence into English (in this test over 500 students out of 843, with only 18 % selecting the correct order). Slightly over half of the items in the test included distractors of this very literal translation type among the possible answers; in 13 items a very large number of candidates selected that alternative (250 - 550) and in a further 9 items a fair number did so (100 or more). What the reasons for this are (the way translation is taught, simple interference from Finnish, etc?) remains to be established, but the phenomenon itself is only too evident, and suggests that more attention needs to be paid to this problem, both on the research level (e.g. the Jyväskylä contrastive study programme) and in practical everyday teaching - in schools and in institutions where English teachers are trained, whether linguistically or pedagogically.

On the basis of this and other tests it should be possible for us to point to specific areas not adequately learnt by students, and to provide a feedback service of information to schools, something which doesn't effectively exist at the moment, and a fruitful dialogue could be commenced.

Glyn Hughes and Anthony May:

ASPECTS OF THE 1974 LISTENING TESTS

1. INTRODUCTION

This paper has a threefold purpose:

(i) to outline the theoretical considerations that inevitably affect the construction of listening tests of this type;

(ii) to review the practical factors that, equally inevitably, applied in the administering of these tests; and

(iii) to isolate some of the problems and issues emerging from the 1974 test results, and on the basis of these make suggestions for improvement or revision in future versions.

Although statistical information will be included where relevant, the primary purpose of the paper is not to present a detailed analysis of the test results.

2. THEORETICAL CONSIDERATIONS

Even in the most seemingly ad hoc assembly of test items, there will be indications that certain principles have informed the test designer, however obscure or undefined these might be. Such principles will have led the designer to produce items which, he honestly believes, will elicit a specific quantity of language behaviour (or probably evidence of such behaviour) from a particular target population, and which will permit him to draw conclusions about the level of language competence demonstrated, and ultimately about the level of language competence acceptable. Thus, a tacit belief in certain methodological tenets coupled with empirical experience of their workability may convince the tester that test format A is the best solution (i.e. the test will be valid and reliable). Similarly, format B can be justified as being the logical implication of a particular psycholinguistic theory that happens to have gained scientific acceptance. Or test format C can have been tried and been seen to work for more or less unexplained reasons.

In reviewing the present set of listening tests, four test situation variables have been identified:

(i) the testees;
(ii) the testing purpose;
(iii) the discourse type; and
(iv) the listening skill.

These variables have to a large extent determined the form of the resulting tests.

2.1. THE TESTEES

Although the candidates taking these tests are products of a more or less standardized educational system, it would be unwise to regard them as homogeneous for testing purposes. There may be tremendous variety in any of the following areas:

(i) the availability and use of audio-visual aids in the candidate's school (i.e. were there any tape recordings available in the school and how effectively were they used?);
(ii) the amount of real language situations that the candidate was obliged to participate in during formal training in the language (was there any actual communicating going on in the classroom, were any real questions asked?);
(iii) the number and type of social encounters that the candidate has experienced outside the classroom where English became a practical medium of communication (has the student met English-speaking people, has he visited Britain or the United States, and in what capacity?);
(iv) the willingness of the candidate to expose himself to the spoken language, and thus put himself in an uncontrolled listening situation;
(v) the candidate's experience of the testing formats used in these tests (multiple choice, note-taking).

Ideally, the imbalance resulting from this diversity of educational and linguistic experience should be eliminated, but is this practicable, or indeed is it our problem at all? Our assumption has had to be that any student seriously considering taking a degree course in English at a Finnish university will require some minimum level of experience of listening to the spoken language, and that on the basis of this experience and further training in situ he should have developed some kind of listening strategy. For this reason the majority of the items in the listening comprehension section of the tests (items 288 - 314) concentrate on isolating a high-level organizational skill rather than range of vocabulary or retention, and thus may function prognostically (see below, 2.2).

It is tentatively suggested that where these skills already exist in the L1 - and they are presumably not unrelated to factors like intelligence[1] - they can be triggered off in the L2 after a certain as yet unidentified period of exposure to the spoken language. It is the job of the schools to provide this exposure.

2.2. THE TESTING PURPOSE

In a listening comprehension test we may be trying to draw conclusions not merely about a candidate's familiarity with elements of the language code (e.g. the range

---

1. Wilkinson and Stratta (1970) report a correlation coefficient of .6036 between an IQ test (verbal and non-verbal) and a wideranging test of listening skills with a sample of 180 13-year-olds. This result is significant at the 1% level.

of his vocabulary), or his retentive powers, but also, and more importantly, about the existence of some overall strategy for listening, some undefined set of 'rules' internalized by the learner that allow him to process speech at sentence, paragraph, discourse level, assign meaning to lexical items on the basis of their context, interpret attitude, personality, adopt an appropriate listening set, etc, etc. In this case, then, we will be testing both attainment and potential. Successful performance on the test items will convince us of the existence of some overall L2 competence, and of the candidate's ability to pursue university studies profitably. For example, the beginning student will be asked to attend 45-minute lectures on unfamiliar topics and to process information in a similar way to that required in the lecturette test (items 305 - 314).

For obvious practical reasons, the test battery does not include any direct measure of productive skills. We must therefore provide ourselves with compensatory tests and on the basis of the results hazard a guess as to the testee's spoken and written English. The phoneme discrimination test (items 229 - 278) may well give clues about the testee's ability to produce the corresponding phonemes in isolation, or indeed it may tell us a great deal about his overall pronunciation.[1]

Likewise, the situationalized response test (items 279 - 287), which measures a complex of skills, including speed of syntactic analysis, awareness of lexical appropriacy, knowledge of social situations, etc., may indicate something about the testee's active experience of and contact with the spoken language, both receptively and productively. Once again, a thorough follow-up study of candidates accepted would verify the predictive power of this test.

2.3. THE DISCOURSE TYPE

Although we are attempting to elicit ecidence of a high-level 'abstract' language ability, the type of language we use in, say, testing listening comprehension should reflect as many aspects of this ability as possible. Ideally, then, we should use the multiplicity of language types that are normally listened to. This does not necessarily mean unscripted language or badly organized language, since there are specific varieties of spoken English demanding careful preparation and delivery. (The "juhlapuhe" is still essentially a spoken thing!) What we should aim at is the widest possible selection of discourse types associated with

---

1. Students accepted to study at Jyväskylä are given a series of active pronunciation tests on entering the English Department. When these tests are judged to be sufficiently reliable, it will be possible to say how far the phoneme discrimination test functions prognostically. See also below, 4.1.

the spoken medium. This selection may be made with certain key variables in mind, for example, the number of speakers, their sex, age, relationship, accent, the particular communicative act they are engaged in (informing, threatening, flirting), the topic of conversation and its associated register features, or the type of listening set presumed. (For a full categorization of these factors, see Chiu, 1973).

The present listening comprehension tests made use of ten different native speakers (7 male, 3 female), and showed a great range in degree of formality, topic, accent, and listening set required. Two of the passages were extracts from unscripted interviews while others were either specially written for testing purposes or extracted from plays and other sources.

2.4. THE LISTENING SKILL

The fundamental distinction made by Brown and Carlsen (1953) between receptive and reflective listening has provided important guidelines in the construction of the test items. An acceptance of the validity of this distinction will lead us to believe that a testee's ability to recognize the overall organizational elements in a piece of spoken English (e.g. item 304), or to identify major details in a passage in preference to amusing subordinate details (item 301), or to draw justifiable inferences (item 289) tells us more about his overall listening skills, and hence his language competence and potential in general than, say, his ability to remember a single detail (a name, a number, an event) from a passage of undifferentiated details.[1]

In the lecturette, too, no attempt is made to test the memory of the candidate. Instead he is required to process incoming information (sift, grade) and encode it in the form of brief notes. Finally he answers a series of multiple-choice questions on the basis of the notes (see below, 4.3).

3. PRACTICAL CONSIDERATIONS

Because of the large numbers of candidates involved, the listening tests must be administered in lecture theatres or examination rooms rather than in language laboratories if all candidates are to complete the test battery on the same day. The sound is amplified and fed to loud-speakers which, if inadequate in number and/or power may favour candidates sitting in the vicinity of the speakers and discriminate against candidates sitting towards the rear of the hall. By moving around the examination hall during answering pauses, we were able to conclude that at Jyväskylä there was no difference detectable by ear in either the quality or the volume of the sound in any

---

1. For sample items see Appendix 6.

part of the hall, except in the immediate vicinity of the loudspeakers (which had, in any case, been kept clear of candidates), where the more noticeable tape hiss made this a less favourable listening area.

3.1. THE 1974 LISTENING TESTS

A master tape was prepared at Jyväskylä and copied, copies being sent to Joensuu and Turku. For all candidates in all three examination centres the listening tests were identical and were administered under similar conditions.

In preparing the 1974 listening tests we felt obliged to consider the following factors:

(1) when they come to the listening tests, the candidates will already have completed the major part of the test battery, with a time allowance of 4 hours; therefore, the growing mental and physical tiredness of the candidates places certain limits on the listening tests, during which the candidates will be under even greater strain than in the earlier tests by virtue of having to concentrate their maximum attention on listening attentively to electronically-reproduced speech;
(2) a sound discrimination test requires a considerable number of items for it to be a valid test;
(3) listening tests that sound artificial, i.e. that are read by no more than one or two RP speakers, may discriminate against those candidates who are less familiar with that particular type of English than with another;

Consideration of these and theoretical factors (see 2. above) led us to adopt the following solutions:

(1) the length of each individual test was kept to a minimum, the longest being 17.5 minutes (including answering time); the different tests were separated from each other by short intervals during which soft music, presumably unfamiliar to most, if not all, candidates was played; the total duration of the listening tests was approximately 55 minutes, during which there were 86 items;
(2) the sound discrimination test was administered in two parts, separated by an interval: Part I consisted of words (minimal pairs) and Part II consisted of sentences;
(3) the dialogues and passages in the listening comprehension test (Test No. 11) were kept as brief as possible; in the lecturette (Test No. 12) candidates were allowed and encouraged to make notes while they listened to the passage.

3.2. SUGGESTIONS FOR THE FUTURE

Two further considerations should be taken into account in designing future entrance examination listening tests: the nature of the listening situation and the choice of passages.

The entrance examination listening situation is extremely artificial: candidates listen with greater than usual concentration to passages reproduced electronically (with the disadvantages of having no visual clues to meaning); relatively large numbers of candidates listen to a series of passages for the purpose of answering multiple choice questions, which they have previously familiarised themselves with. Perhaps the only activity which this process resembles is performing some of the tasks of listening comprehension exercises, not in itself a major part of students' university studies.

The choice of passages, too, may be described as, for want of a more accurate word, arbitrary. The test consists of items which, in general, discriminate well, unsuccessful items having been rejected in previous years. However, the test cannot be described either as an achievement test nor, in its present format, as a prognostic test.

It should be possible to meet these objections by designing listening tests which reflect more accurately the candidates' pronunciation and oral production skills and place the candidates in a more realistic listening situation. Proposals for such tests will be discussed in the following section.

4. THE INDIVIDUAL TESTS

4.1. THE 1974 SOUND DISCRIMINATION TEST

The test was administered in two parts: Part I consisted of 30 items (words), Part II consisted of 20 items (sentences). In each item candidates were asked to identify which one of the three words or short sentences differed from the other two. The phonemes around which this test was designed are those which cause most difficulty to Finns when they pronounce English (in particular plosives, fricatives, affricates). For reasons already expressed (see 2.1. & 2.2.) the test cannot act as an achievement test; it would be reasonable to regard it as a prognostic test, but, in the absence of any follow-up study of successful candidates' performance, there are no valid criteria for judging it accordingly. It would, however, seem reasonable to expect the results of an effective sound discrimination test to be roughly similar to those of other tests in the battery. In 1974 this was not exactly the case.

The results show that the Sound Discrimination Test contained 13 items with a solving percentage in excess of 90; all of these items were in Part I, representing more than 40% of that sub-test. An example of this category is given below:

Item 236 a) shoe b) Sue c) shoe
Result: b) r = .58, solution percentage 94.07
a) r = -.09, N = 25
c) r = -.13, N = 6.

Such high solution percentages indicate that the test has lost its effectiveness as part of a test battery designed for selecting fewer than 25% of the total number of candidates.

If one examines the results of the two sub-tests, certain differences come to light: in Part I the average solution percentage was 86.82, the range being from 67.97 to 97.15; in Part II, however, the average solution percentage was 70.20, the range being from 43.42 to 89.32. Two very similar items, one from Part I, the other from Part II, exemplify these differences.

Item 250 a) age b) age c) aids
Result: c) r = .55, solution percentage 84.19
a) r = -.06, N = 23
b) r = -.07, N = 7

Item 270 a) We had no defence against his rage.
b) We had no defence against his raids.
c) We had no defence against his raids.
Result: a) r = .16, solution percentage 43.42
b) r = -.03, N = 228
c) r = -.12, N = 228.

As it is most unlikely that, within the space of a few minutes, 50.77% of the candidates lost their ability to distinguish between dʒ and dz, one must look elsewhere for an explanation. In the first subtest the candidate is listening to three words out of context; he is, thus, concerned with sounds and not with meaning. In the second subtest the candidate is given a context, the sentence, and may, therefore, automatically try to attach a meaning to the whole sentence, including the words concerned; his power of concentration on the minimal pair is thereby weakened, a memory element being introduced, as the three words are now separated by the intervening sentence. If this is, indeed, the case, using sentences in this way is not as effective a way of testing sound discrimination as is using single words.

4.1.2. THE 1973 AND 1974 SOUND DISCRIMINATION TESTS COMPARED: SOME FACTORS AFFECTING DIFFICULTY

Not only were the solution percentages for the 1974 Sound Discrimination Test higher than for the other tests in the battery, they were also higher than in the 1973 test:
1973, 62.01% 1974, 80.17%.
In spite of the 'better' results in 1974, tentative judgments at Jyväskylä based on the pronunciation pre-test given to all new Approbatur students, suggest that, as far as pronunciation is concerned, the 1974 intake is no better than the 1973 intake.

One may, however, suggest an explanation for the above test result differences. In 1973 the test consisted

of 3-word and 3-sentence items (as did the 1974 test) but had a higher task complexity: candidates were asked to indicate which one of the three words/sentences differed from the other two or, by marking a fourth alternative, to indicate that all three words/sentences were different, which was the case in some items. It is, therefore, possible to criticise the 1973 test for placing a memory load on candidates who are already in a stress condition. The 1974 test had a reduced number of alternatives, three instead of four (the alternative indicating that all three words/sentences were different being eliminated); with the resulting reduced task complexity, the memory factor was also reduced. But, as the candidates knew in the 1974 test in advance of each item that one word/sentence was going to be different, they may have been encouraged to listen not so much to the sounds themselves but, rather, to the differences between the sounds. The candidate might, thus, have selected the correct alternative but have still been unable to identify the same word if, having heard it, he had then been asked to state which one of a list of words on his question paper he had heard.

The reduced task complexity may not, alone, explain the differences in results between the two years. In 1974 the readers of both subtests were different to the readers of the corresponding subtests in 1973. (See 4.2.2. for a further discussion on this point.)

4.1.3. SUGGESTIONS FOR AN IMPROVED SOUND DISCRIMINATION TEST

It might be worth considering the following proposal for the 1975 entrance examination. Each item would consist of a single word/sentence which is repeated two or more times; the candidate would be required to select from the three or four alternatives on his question paper the word he thinks he has heard, as in the following example:

Tape: peas - peas - peas

Question paper: (a) bees (b) peas (c) peace (d) piss

This procedure would reduce the memory factor even further and would ensure that the candidate listens to and identifies sounds rather than the differences between sounds; the present phoneme discrimination test would be transformed into a phoneme identification test, which need not be based on the minimal pair principle.

4.2. TEST NO. 11: SITUATIONALISED RESPONSES AND LISTENING COMPREHENSION

Although this test is described, in the statistical analysis and elsewhere, as 'Listening Comprehension', in fact it consisted of two sub-tests: 1) Situationalised Responses, comprising 9 items; 2) Listening Comprehension, comprising 17 items, 8 of which were short dialogues, the remaining 9 being short passages of various kinds, one including a short dialogue, two being extracts from unscripted interviews, and the remainder being

classifiable as, variously, reminiscence, expression of opinion about a third person, anecdote, description of actions, exposition.

4.2.1. SITUATIONALISED RESPONSES

In a test battery of the type described in the present paper, there is no possibility for measuring the candidates' own oral production, nor would there be any possibility of attempting to do so objectively without considerably increasing the time required for the test battery and dalaying the availability of results.

The closest the present battery comes to estimating the candidates' ability to speak reasonably accurate English is by means of the situationalised responses sub-test, where the candidate hears an everyday question or request and then selects the appropriate response from among the alternatives on his question paper, as in the following example:

Item 282 Tape: 'Would you like to go for a drive in the car today?'
a) Yes, I do.
b) Yes, I would.
c) Yes, I like to.

Result: b) r = .25, solution percentage 70.58
a) r = -.17, N = 37
c) r = -.18, N = 192.

It may be objected that a sub-test of 9 items is insufficient, and it should be possible to meet this objection by re-designing this sub-test, perhaps by including a number of items where the appropriate completion of incomplete spoken passages must be selected. The resulting test might well consist of at least 20 items.

4.2.2. EXTERNAL FACTORS AFFECTING DIFFICULTY: A COMPARISON OF THE 1973 AND 1974 TESTS

All the items in the 1974 test had also featured in the 1973 test. They were retained substantially unchanged, but, except for two passages extracted from unscripted interviews, all the passages were recorded anew for the 1974 test with, in most cases, different voices to those employed in 1973. Four items that had proved to be poor, according to the computer analysis, were rejected altogether.

In the two identical items extracted from unscripted interviews the solution percentages were very similar in 1973 and 1974:

	1973	1974
Item A:	69.60	69.28
Item B:	55.50	58.01

Furthermore, in two short dialogue passages both readers were the same in 1973 and 1974; one of them is given below as an example:

Tape: Mr. Smith: Hello, James. What can I do for you?
James: Would you mind if I had some time off, Mr. Smith?
Mr. Smith: When, exactly?
James: Monday and Tuesday of next week.
Mr. Smith: I'd like to say yes, but it's just not possible.

Item 295 Mr. Smith -
a) will let James do what he wants to do.
b) will give James an answer the following Monday or Tuesday.
c) cannot let James do what he wants to do.

Result: 1973
c) r = .26, solution % 91.33
a) r = -.17, N = 14
b) r = -.36, N = 53
1974
c) r = .48, solution % 91.10
a) r = -.22, N = 16
b) r = -.23, N = 38

In the other item where the readers were the same in both years the relevant statistics were as follows:

Item 294
Result: 1973
b) r = .28, solution % 68.43
a) r = -.07, N = 8
c) r = -.49, N = 236
1974
b) r = .35, solution % 71.06
a) r = -.12, N = 18
c) r = -.30, N = 207

One item, in spite of being rejected by the computer in 1973 on account of its negative item correlation, was retained in 1974 in order to test the hypothesis that a different reader and a different reading could affect the result.

Tape: Bill and Ann are talking to James.
Bill: We met Mr. and Mrs. Smith yesterday, James. He told us that they were thinking of buying a new house.
Ann: She told us, Bill.

Item 292
a) Mr. Smith told Bill and Ann something.
b) Mrs. Smith told Bill and Ann something.
c) Both Mr. and Mrs. Smith told Bill and Ann something.

Result: 1973
b)+r = -.04, solution % 46.96
a) r = -.19, N = 165
c) r = -.16, N = 245
1974
b) r = .22, solution % 62.99
a) r = -.23, N = 45
c) r = -.14, N = 247

The part of Bill was read by the same person in both years, but the part of Ann was read by different people in 1973 and 1974, and, furthermore, especial care was taken in 1974 over the way in which Ann's part was read. It seems justifiable, therefore, to attribute part of the significant difference in the results for the two years to reading differences (voice, intonation, emphasis).

Other factors affecting difficulty should be considered under the heading 'external factors', the first being the choice of language for the questions. The passages themselves were of no more than moderate difficulty, and the questions, in English, were worded in as simple terms as possible. All the same, it may be objected that there is an element of reading comprehension in the test, but if the questions were phrased in the candidates' native language, it might then be objected that the reading comprehension element has merely been replaced by a translation element, as the candidates would then be required to correlate L2 passages and L1 questions; at least some candidates might find this confusing (as used to be the case in some previous Jyväskylä entrance examinations).

There is a further argument in favour of having questions in English: especially in a multi-lingual or bi-lingual country, such as Finland, not every candidate can automatically be assumed to be a native speaker of the majority language, even if they may be expected to understand instructions in that language. Questions in English have the merit of not giving any non-Finnish-speaking candidate reasonable grounds for feeling discriminated against because of his native language.

A further external factor which might affect difficulty is whether the questions should be read before or only after the relevant passage has been heard. The argument in favour of the latter is that the candidates are then required to concentrate on the passage as a whole and their understanding of the passage can provide the basis for answering the question; they would have no previous clues as to what they should concentrate on while listening, a procedure which, it might be thought, more closely resembles normal listening situations in real life than the other, more artificial, procedure. The argument in favour of the former (letting candidates read the questions before hearing the passage) is that this reduces the memory factor and enables the candidates to focus their attention on the passage more effectively, though with the concomitant risk that some candidates may be encouraged thereby to establish an unwarranted connection between a distractor on their question paper and the part of the passage to which this distractor refers. It is, however, administratively easier to let candidates read the question before hearing the relevant passage, as has been the case in 1973 and 1974.

4.2.3. INTERNAL FACTORS AFFECTING DIFFICULTY

The mean solution percentage for this test was 72.95, making it one of the easier sections of the battery. Some individual items, however, stand out as being considerably more difficult, and it is these that will now be examined. Consider, for example, the following item:

> He decided to watch what was happening. Arnold was just getting into the ring and the crowd naturally gave him a big hand. He clambered through the ropes and put his two fists together and gave the crowd a big smile, shaking his fists all the while, confident as a real champ should be. Then he sat down as Evans came down the passageway through the crowd. Evans was a Welshman, so he didn't get such a big hand as the London boy, but he smiled nevertheless, with one gloved fist in the air. He climbed up and bent down to get through the ropes. Arnold saw this and got up from his corner to push the rope down so that Evans could get through. The crowd thought that this was wonderful. Arnold put his hand on Evans's shoulder and they stood there for just a moment.

Item 303 Arnold and Evans are about to -

a) climb a mountain
b) do a circus act
c) start a boxing match

Result:

c) r = .30, solution percentage 55.52
a) r = -.22, N = 112
b) r = -.20, N = 228

1973 result:

c) r = .20, solution percentage 53.56
a) r = -.35, N = 114
b) r = -.27, N = 243

This is a good item with both distractors working well. The comparative difficulty of the item lies in the nature of the listening task set the candidate. How does the candidate tackle this problem? Presumably he initially reads the question and alternatives. As the tape begins, he is seeking confirmation for one of three already existing hypotheses; in other words, he is being asked to focus his listening on items that suggest a circus act, mountain climbing, or boxing. But simultaneously, he must correlate these lexical clues, and reduce the multiplicity of real world referents to one specific contextual field. Many of the words (e.g. rope, climb, crowd, ring) point in several directions simultaneously. The native speaker is in the same position when he enters a conversation half way through and must listen retrospectively. Behaviour typical of the boxing ring is indicated repeatedly in the passage, but unless the unifying context is discovered, the behaviour becomes meaningless and is assigned to the circus arena, where illogical behaviour is, as it were, contextualized.

Candidates failing this item were incapable of making use of lexical clues. This indicates a paucity of experience of the real language, and also reflects competence deficiencies.

> I knew Ralph Vaughan Williams well when he was an undergraduate at Cambridge, though not as well as I knew his cousin, Ralph Wedgwood. I was very fond of Ralph Vaughan Williams, who was in those days a most determined atheist and was noted for having walked into hall one evening saying in a loud voice, 'Who believes in God nowadays, I should like to know?' I hardly ever saw him in later years, though I retained an affectionate admiration for him.

Item 299 The anecdote told by this speaker indicates that at Cambridge Ralph Vaughan Williams -

a) believed in God
b) did not believe in God
c) was trying to find out how many people believed in God.

Result:

b) r = .30, solution percentage 52.43
a) r = -.11, N = 40
c) r = -.29, N = 326

1973 result:

b) r = .25, solution percentage 46.57
a) r = -.24, N = 64
c) r = -.40, N = 347

Distractor a) only misled a small number of candidates, suggesting generally that the word 'atheist' was understood. The difficulty here lies in the failure of a large number of candidates to understand the communicative function of Williams's question, 'Who believes in God nowadays, I should like to know?' In this context it can only be interpreted as a direct challenge, and clearly the superficial request form (I should like) for information (to know) does not presuppose an equivalent function semantically. Where a candidate's listening is confined to the receptive level, this item will present considerable difficulty.

From the evidence of some of the easier items, it seems that many of the candidates do possess an inferential listening skill in the L2. However, this skill would appear to be restricted to recurrent everyday contexts where attitude and intention may be expressed phraseologically; in other words, the inference is embodied in the particular choice of phrase. For example:

Mr. Smith: Hello, James. What can I do for you?
James: Would you mind if I had some time off, Mr. Smith?
Mr. Smith: When exactly?
James: Monday and Tuesday of next week.
Mr. Smith: I'd like to say yes, but it's just not possible.

Item 295 Mr. Smith -
a) will let James do what he wants to do
b) will give James an answer the following Monday or Tuesday
c) cannot let James do what he wants to do

Result:
c) r = .48, solution percentage 91.1
a) r = -.11, N = 16
b) r = -.23, N = 38

1973 result:
c) r = .26, solution percentage 91.33
a) r = -.17, N = 14
b) r = -.36, N = 53

4.2.4. SUGGESTIONS FOR REVISED TEST FORMAT

Whereas items 303 and 299 test discourse skills that belong well and truly to reflective listening, the simpler items, like 295, may in fact be more closely related to some kind of L2 communicative competence, and could perhaps be tested more effectively in a modified and extended version of the situationalized responses test. (The existing test consists of a mere nine items.) Such a revision would allow simultaneous testing of receptive listening and awareness of grammatical, lexical, and situational appropriacy. Two possible formats for such a test are suggested below.

Format 1: Tape:

You are travelling by train from London to Manchester. A young man turns a radio on and you find it impossible to sleep. You tell him to turn it off. Not only is it annoying you, but it is against the law. He apologizes and says: 'I didn't know that playing a radio was forbidden in trains.'
What do you now say?

Answer sheet:
a) Nor did I
b) It isn't
c) It is though
d) I'm very sorry about that

Format 2: Tape:

'A. Well, I've thought about it, and I know I ought to go and see it. It is after all one of his best films and it's only on for three days. Anyway, the fact of the matter is I just can't wait to see it.'
'B. Well, I've thought about it, and I know I ought to go and see it. It is after all one of his best films and it's only on for three days. Anyway, the fact of the matter is I thought it was a waste of time.'
'C. Well, I've thought about it, and I know I ought to go and see it. It is after all one of his best films and it's only on for three days. Anyway, the fact of the matter is I just haven't got the time to spare.'

Answer sheet: Which speaker made sense?
a) A, b) B, c) C, (and possibly: d) none of them, e) all of them)

4.3. THE LECTURETTE

This was found to be one of the most difficult sub-tests (mean solution percentage, 56.81%). The more difficult items from the listening comprehension test examined above (4.3.) had solution percentages of a comparable level (item 303, 55.52%; item 299, 52.43%; item 301, 58.01%). This suggests that the same kind of higher level listening strategy is being isolated. An additional complicating factor is probably the unfamiliarity of the lecture situation and the vagueness of the instructions given to the candidates: 'Tee luentoa kuunnellessasi lyhyitä muistiinpanoja keskeisistä asioista.' (While listening to the lecture, make brief notes on the main ideas.) The lecture is also the final sub-test and comes after several hours of concentrated labour.

What the candidate should be doing is grading incoming information and organizing it along particular lines, but this has to be done extremely rapidly. So, for example, the definition of linguistic elevation is given and exemplified briefly in the lecture, and the student has the following to answer:

Item 312

Which of the following words has undergone elevation in the last two or three centuries? (The 17th or 18th century definition is given in Finnish.)

a) Apparatus = varusteet, tarvikkeet
b) Formula = vakiintunut sanamuoto, jolla jokin asia määritellään, todetaan tai selitetään
c) Genius = synnynnäinen luonteenlaatu, luonnollinen taipumus

Result.

c) r = .24, solution percentage 49.70
a) r = -.39, N = 225
b) r = -.34, N = 169

1973 result:
c) r = .13, solution percentage 49.03
a) r = -.42, N = 229
b) r = -.40, N = 155

The trouble here is that the student is not given sufficient time to process the information. A successful student may have deduced the correct answer purely from the written question, without reference to his notes or the lecture. An improved version of this test would have to include the normal 'built-in' redundancy habits of the university lecturer (repetition, recapitulation, exemplification, emphasisingof central points, use of drawings and OHP transparencies, ets.). Also performance variables (false-starts, hesitation, coughs) and

paralinguistic clues (gesture, facial expression) belong inseparably to the lecture situation. Given these additional features, note-taking becomes feasible.

4.3.1. SUGGESTIONS FOR REVISED TEST FORMAT

An improvement affecting the whole battery would be the redistribution of the listening comprehension items over two tests. One section would contain situationalized responses (see above, 4.3.) - in other words, concentrating on receptive listening and communicative competence - whilst the lecture would be expanded to 15 - 20 minutes' duration and would be presented either via film or video-tape recording. Alternatively, the lecture could be replaced by a round-table discussion, thus allowing factors such as accent, personality, and attitude to be included. The video-tape presentation is presumably not impossible in practical terms, and all the existing defects of the lecturette test could be removed (unnaturalness, no redundancy, etc.).
Even in this revised form, it would allow testing of reflective listening skills and would serve a valuable prognostic function.

REFERENCES

Brown, J.I. and Carlsen, G.R. (1953). Brown-Carlsen Listening Comprehension Test, Evaluation and Adjustment Series, Grades 9-13. World Book Co., Yonkers, N.Y.

Chiu, R.K. (1973). 'Measuring Register Characteristics', IRAL, XI, 1.

Wilkinson, A.M. and Stratta, L. (1970). 'Listening Comprehension at 13-plus'. Educational Review, 22.3.

Väinö Roininen - Pirkko Kosonen:

THE 1974 JOINT STUDENT SELECTION FOR ENGLISH: CANDIDATES' OPINIONS ABOUT THE INFORMATION, THE TEST BATTERY USED AND THE INTER-UNIVERSITY CO-OPERATION

1. INTRODUCTION

The present paper gives the main results of a study carried out at the Department of Sociology, University of Turku, the aim of which was to find out student opinion on the joint entrance test of the English departments of Joensuu, Jyväskylä and Turku in summer 1974. The study covers the following areas:

a) advance information and test instructions
b) the test battery as an instrument of student selection
c) desirability of a joint entrance test and extended co-operation

The 1974 entrance test was the second jointly planned test of the three departments of English, administered simultaneously in the three places. The whole test battery consisted of 12 tests, of which 7 were taken in the actual entrance examination, the remaining 5 being school grades. The former tests accounted for 70 % and the latter for 30 % of the final scores on the whole test battery, and the selection decisions were based on this final score (for a more detailed analysis of the tests, see Kohonen's paper).

The joint entrance examination represents an attempt to use computer technology to develop a valid, reliable and economical testing instrument, which is also desirable from the candidates' point of view. As entrance tests are usually arranged separately by different departments, candidates generally have to travel to several places, send in several applications, take many copies of their certificates, acquire several books etc., which involves considerable costs and time. In a joint selection, however, such costs will be greatly reduced, and the co-operation between the three universities may be seen as an attempt to rationalize student selection (cf. Parjanen 1972, 2). For the first time, in the present selection project, a questionnaire was given to a sample of the candidates for the purpose of obtaining information about student opinion needed for the further development of the selection procedures. In addition to the manifest function of student selection, entrance tests also have latent functions, such as the formation of attitudes towards the universities in question as the result of the entrance examination. If it is felt that the selection procedures

are fair, the attitudes are likely to be positive, which may affect students' later adjustment and studying motivation (cf. Carlsson 1965, 25-26). This is another reason why it is important to know student opinion on selection. One could say that this is a question of occupational democracy: the students are being asked to give their opinions on important questions concerning their work.

2. POPULATION AND RESEARCH METHOD

The population of the study consisted of the 843 candidates for the universities of Joensuu, Jyväskylä and Turku. The data was collected using a questionnaire[1] (see Appendix 3, in Finnish) which was given to every third candidate (i.e. every third row of seats in the lecture hall, which actually meant 41.2 per cent of the population) at the end of the examination at each of the universities. Consequently, the sample consisted of the following numbers:

Joensuu	46	candidates
Jyväskylä	191	"
Turku	110	"
Total	347	candidates

It is assumed that the sample is representative. For instance, the sample mean score does not differ statistically from the population mean:

	N	$\bar{X}$	standard error
population	843	222.67	30.98
sample	347	224.52	33.40

3. OPINIONS ON THE ENTRANCE TEST PROCEDURES

3.1. Advance information

Candidates were given advance information about the selection procedures in the printed study guides published by each university (the Turku information is given in Appendix 1 and that for Jyväskylä in Appendix 4, in Finnish). At the beginning of the examination, every candidate received a duplicated copy of the test instructions (presented in Appendix 2), with a practice example, which was carried out and explained on the blackboard by the supervising teacher. In addition, there were brief instructions at the beginning of each test, and in the listening tests these were also recorded on the tape. The adequacy of this information is examined in Tables 1 and 2.

---

[1] The authors are indebted to Lilja and Viljo Kohonen for help in the planning of the questionnaire and for practical help in data processing, as well as for translating this paper.

Table 1. Candidates' opinion on the adequacy of the advance information, by sex (M = male, F = female, T = total) and place of examination (%).

Adequacy of advance information	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
Entirely inadequate	12	-	4	17	20	19	15	8	9	16	14	14
Reasonable, but not adequate	25	43	37	39	40	40	30	34	34	34	38	37
Fairly adequate	38	43	41	26	26	26	40	40	40	31	33	32
Completely adequate	25	13	17	17	13	14	15	18	17	18	15	16
Total	100	100	100	100	100	100	100	100	100	100	100	100
	(16)	(30)	(46)	(46)	(142)	(188)	(20)	(90)	(110)	(82)	(262)	(344)

Half of the candidates indicated that advance information about the entrance examination was somewhat inadequate. The opinions do not depend on sex, but there are differences according to where the test was taken: those who applied for Jyväskylä considered the advance information inadequate more often than the others. The reason for this is probably that the advance information given at Jyväskylä was rather scanty. The information thus does not seem to have been adequate, and should be improved in future in order to set the candidates on an equal footing. According to Table 2, relatively least information was

Table 2. Demand for advance information about the different tests in the three testing places.

Test	Joensuu	Jyväskylä	Turku	Total
Vocabulary and Grammar I	15	18	13	16
Vocabulary and Grammar II	19	21	22	21
Translation Analysis	15	18	15	16
Reading Comprehension	17	26	14	21
Auditory Discrimination	15	29	38	30
Listening Comprehension	24	21	21	21
Lecture	26	39	40	38
	(N=46)	(N=191)	(N=110)	(N=347)

given about the lecture and the auditory discrimination tests; the request for more information about the lecture was greatest in Jyväskylä and Turku, and for the auditory discrimination in Turku.

Advance information is also a question of fairness. Candidates should be given sufficient information about the nature of the tests, their relative weights in the selection, and the procedure for applying for a revised selection decision in case of injustice (Korkeakouluhallinnon oikeusturvakomitean mietintö 1971: A4, 51). If sufficient information is not provided to all applicants, The student selection may be affected by irrelevant factors. It has been found, for example, that parents who have themselves received higher education can inform their children about the requirements of university studies better than those from lower socio-economic classes, and thus give more help to their children (Marin 1970, 149).

The present entrance examination has inadequacies in all these areas and in all the three testing places, most of all, however, in Jyväskylä, where the printed study guide gives no information about the nature of the entrance examination and has, in addition, ambiguities that may have been confusing.

3.2. Test instructions

Student opinion on the test instructions is summarized in Table 3.

Table 3. Clarity of the test instructions, by sex and place of examination (%).

Clarity of instructions	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
Partly unclear	-	-	-	9	9	9	5	7	6	6	7	7
Quite clear	19	30	26	35	36	36	20	26	24	28	32	31
Entirely clear and understandable	81	70	74	55	55	55	75	67	68	66	61	62
Total	100	100	100	100	100	100	100	100	100	100	100	100
	(16)	(30)	(46)	(47)	(144)	(191)	(20)	(90)	(110)	(83)	(264)	(347)

Most of the candidates, irrespective of sex and testing place, considered that the test instructions were clear and understandable. Only 7 per cent of the candidates thought that the instructions were partly unclear. Those who took the test in Jyväskylä had different opinions from the others: they considered the instructions "quite

clear" more often than the others, and "entirely clear and understandable" less frequently than the others.

Table 4. Percentages of candidates who indicated that the test instructions should be improved, by sub-tests and testing place.

Test	Joensuu	Jyväskylä	Turku	Total
Vocabulary and Grammar I	7	2	3	3
Vocabulary and Grammar II	4	2	7	4
Translation Analysis	4	8	5	6
Reading Comprehension	15	24	16	20
Auditory Discrimination	2	5	7	5
Listening Comprehension	-	2	5	3
Lecture	11	8	6	8
	(N=46)	(N=191)	(N=110)	(N=347)

There is relatively most demand for clearer instructions in the Reading Comprehension tests, particularly among the Jyväskylä group. In other tests, the instructions were generally felt to be good, since the percentages of those who wanted better instructions are remarkably low.

3.3. Sufficiency of the time reserved for the tests and the suitability of the examination date

The time reserved for the written tests (total of 223 items) was 4 hours, from 10 a.m. to 2 p.m., and for the listening tests (86 items), 55 minutes, from 3 to 4 p.m.. In the written tests, candidates were free to leave the room after they had finished (which most of them did long before 2 p.m.).

Table 5. Sufficiency of time allocation for the written and the listening tests, by sex and the testing place.

Time allocation and tests	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
Written tests												
Insufficient	-	10	7	-	3	2	-	6	5	-	5	4
Fairly insufficient	25	23	24	15	12	13	40	29	29	23	19	20
Entirely sufficient	75	67	70	85	85	85	60	66	65	77	76	76
Total	100	100	100	100	100	100	100	100	100	100	100	100
	(16)	(30)	(46)	(47)	(143)	(190)	(20)	(90)	(110)	(83)	(263)	(347)

Listening tests												
Insufficient	6	17	13	15	10	11	33	14	16	17	13	14
Fairly insufficient	50	43	46	44	54	51	28	60	55	41	53	50
Entirely sufficient	44	40	41	41	36	37	39	25	28	41	34	36
Total	100	100	100	100	100	100	100	100	100	100	100	100
	(16)	(30)	(46)	(41)	(130)	(171)	(18)	(98)	(116)	(75)	(258)	(333)

Three quarters considered the time allocation for the written tests to be wholly sufficient and only a small number of females felt that it was insufficient; this was particularly the case in the Jyväskylä group. A comparison of the written and listening tests shows that the time allocation for the latter was considered insufficient more often than for the former. The Turku group, particularly, felt that the listening tests should have been given a longer time allocation, and males more often than females. In general, males indicated that the time was "entirely sufficient" more often than females, who felt that the time allocation was "fairly sufficient", particularly in the Turku group.

The examination took place on July 8th, and candidates were asked how suitable this date was by having them check the preference for one of the dates suggested.

Table 6. Suitable date for the entrance examination, according to sex and the testing place.

Date of the examination	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
no answer		7	6		3	3		1	1		3	3
15.6.-30.6.	6	13	9	21	24	23	5	21	18	13	22	18
1.7.-15.7.	31	47	41	51	47	48	35	39	38	43	46	44
16.7.-31.7.	63	20	35	17	17	17	35	22	24	31	20	23
1.8.-15.8.	0	13	9	0	7	6	15	13	14	6	10	10
16.8.-31.8.	0	0	0	6	1	3	10	3	4	6	2	2
Total %	100	100	100	100	100	100	100	100	100	100	100	100

July is regarded as the best suited date for the examination by two thirds of the candidates, June by one fifth and August by one eighth. An earlier date seems generally to be preferred by females.

4. APPROPRIATENESS OF THE TEST BATTERY AS A STUDENT SELECTION INSTRUMENT

4.1. Appropriateness of the test battery

About one third of the candidates was unable (or unwilling) to evaluate the appropriateness of the present test battery or parts of it as a language test and a student selection instrument. This "lack of opinion" was more common in the Jyväskylä group, particularly among the females (Table 7). This may be partly due to the scanty advance information given in Jyväskylä. On the other hand, such a central tendency in the answers is a general feature in sociological studies (Eskola 1968, 152). Evaluations may also be influenced by such factors as tiredness (the questionnaire was completed after the 5-hour examination), and differences in attitude due to geographical situation; these cannot, however, be considered here.

Table 7. Student opinion on the appropriateness of the test battery as a student selection instrument, by sex and testing place.

Test and testing place	Appropriateness 1	2	3	Appropriateness 1	2	3	Appropriateness 1	2	3
Vocabulary and Grammar I	7	24	69	8	24	68	8	24	68
Joensuu	0	31	69	10	23	67	6	26	68
Jyväskylä	9	21	71	7	28	66	8	26	67
Turku	10	25	65	9	18	74	9	19	72
Vocabulary and Grammar II	13	25	61	10	28	63	10	27	63
Joensuu	6	25	68	17	23	60	13	24	63
Jyväskylä	15	25	60	7	31	62	9	30	60
Turku	15	25	60	10	23	67	11	24	66
Translation Analysis	7	33	60	11	30	59	10	31	59
Joensuu	0	50	50	10	26	63	7	35	59
Jyväskylä	9	36	55	10	34	56	10	35	55
Turku	10	10	80	12	24	64	12	22	66
Reading Comprehension	13	29	58	16	35	49	15	33	52
Joensuu	0	44	56	20	20	60	13	28	59
Jyväskylä	17	34	49	16	44	40	17	41	43
Turku	15	5	80	13	25	61	14	22	65
Auditory Discrimination	26	29	44	20	30	50	22	29	50
Joensuu	12	36	51	23	17	40	20	24	56
Jyväskylä	23	30	47	17	36	47	19	34	47
Turku	45	20	35	24	23	53	27	23	50
Listening Comprehension	17	29	54	10	30	60	11	30	60
Joensuu	12	31	57	3	20	77	6	24	69
Jyväskylä	15	28	58	11	38	51	12	35	54
Turku	25	30	45	9	21	70	11	23	65

Lecture	22	25	53	19	37	45	19	34	46
Joensuu	25	37	38	17	13	70	19	22	59
Jyväskylä	21	21	58	18	47	36	19	40	41
Turku	20	25	55	20	30	50	20	29	51
Test as a whole	6	25	69	5	37	58	6	34	60
Joensuu	0	25	75	3	30	67	2	28	70
Jyväskylä	6	23	71	5	43	52	5	38	57
Turku	10	30	60	7	31	62	7	31	62

1 = poor or very poor
2 = uncertain, cannot say
3 = good or very good

Three fifths of the candidates indicated that the whole battery was a suitable student selection instrument, and only 6 per cent said that it was poor. The appropriateness of the lecture and the auditory discrimination tests was questioned by one fifth, reading comprehension by one seventh, and the remaining tests by about one tenth. Three fifths considered Vocabulary and Grammar I to be an appropriate test, and a similar opinion about the auditory discrimination, lecture and reading comprehension tests was indicated by half of the candidates.

The evaluations show some differences that depend on sex and the testing place. Males indicated, more frequently than females, that auditory discrimination and listening comprehension were poorly suited, and this was particularly the case in Turku. The whole test battery was found acceptable by the males more often than by the females, and in the Joensuu and particularly the Jyväskylä groups, the males have the most positive evaluations. In the Turku group, the males indicated more often than the females that translation analysis and reading comprehension were appropriate tests.

It can be concluded that most of the candidates regard the test battery as appropriate and consider that the selection of students on the basis of its results is fair.

4.2. Weight coefficient of the entrance test and the school grades

In the present selection, the total weight coefficient of the different entrance tests was 70 % and of the school grades 30 %. Candidates were asked to give their opinion on these proportions.

Two thirds of the candidates consider that the present proportion 70 % and 30 % is suitable, the males more frequently than the females. One third think that the weight of the entrance test should be reduced, and one tenth think that it should be increased. These opinions depend on sex and the testing place, so that the females

Table 8. Student opinion on the weight coefficient of the entrance test and the school grades, by sex and the testing place.

Suggested proportional weight for the entrance test	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
Below 50 %	12	0	4	2	7	6	5	3	3	4	5	5
50 - 59 %	6	7	7	15	24	22	5	30	26	11	24	21
60 - 69 %	6	13	11	2	7	6	10	4	6	5	7	6
70 - 79 %	56	60	59	62	56	59	60	57	57	66	57	58
80 - 89 %	19	10	13	2	1	1	10	3	4	7	3	4
90 -100 %	0	10	7	17	4	7	10	2	4	12	4	6
Total %	100	100	100	100	100	100	100	100	100	100	100	100
	16	30	46	47	144	191	20	90	110	83	264	347

in the Jyväskylä and Turku groups preferred a lower weight coefficient for the entrance test and were less willing to increase it. One fifth of all males also indicated their preference for a lower weight coefficient, while again one fifth of them wanted to increase it.

The present proportion of 30 to 70 is probably good; the study by Kivekäs and Parjanen (1971, 24) suggests that the school grades should not be given much weight, since their predictive ability in university studies seems quite low.

4.3. The present examination versus a translation test

Candidates were also asked to indicate which type of a test they would prefer in the student selection, i.e. between a translation test (English-Finnish and Finnish-English), known to everybody from the secondary school, or the type of test used in 1974, i.e. a multiplechoice test. The results are given in table 9.

Table 9. Student opinion on the type of examination, whether a multiple choice test or a translation test, by sex and testing place.

Opinion and testing place	Male	Female	Total
Joensuu			
Translation test better	12	23	20
Cannot say	19	7	11
This test better	69	70	70
Total	100	100	100
	(16)	(30)	(46)

Jyväskylä			
Translation test better	21	12	14
Cannot say	13	19	17
This test better	66	69	69
Total	100	100	100
	(47)	(144)	(191)
Turku			
Translation test better	20	22	22
Cannot say	10	10	10
This test better	70	68	68
Total	100	100	100
	(20)	(90)	(110)
Total			
Translation test better	19	17	17
Cannot say	13	14	14
This test better	68	69	69
Total	100	100	100
	(83)	(264)	(347)

Two thirds of the candidates indicated their preference for the present kind of multiple-choice test, while one fifth said that they would have preferred a translation test; these opinions were independent of sex and testing place.

4.4. Level of difficulty of the entrance examination

As it is interesting to know how difficult or easy the test was in the candidates' opinion, there was also a question about this, where the assessments were given using a scale from 1 (easy) to 4 (very difficult).

Table 10. Opinion on the difficulty level, according to sex and the testing place.

Difficulty level and testing place	Male	Female	Total
Joensuu			
Easy	6	20	15
Suitable	38	43	41
Difficult	44	30	35
Very difficult	12	7	9
Total	100	100	100
	(16)	(30)	(46)
Jyväskylä			
No answer		5	3
Easy	13	23	20
Suitable	34	40	38
Difficult	40	30	33
Very difficult	13	4	6
Total	100	100	100
	(47)	(144)	(191)

Turku			
No answer		3	3
Easy	10	19	17
Suitable	20	40	36
Difficult	65	32	38
Very difficult	5	5	5
Total	100	100	100
	(20)	(90)	(110)
Total			
No answer		3	2
Easy	11	22	19
Suitable	31	41	38
Difficult	47	32	35
Very difficult	11	5	6
Total	100	100	100
	(83)	(264)	(347)

One fifth felt the examination was easy, two fifths indicated that the difficulty level was suitable, and two fifths said that it was difficult. The males generally considered it to be more difficult than the females. Female opinion was quite unanimous in all the three places, while the males in the Turku group regarded it as difficult relatively more frequently than did the males in Jyväskylä and Joensuu.

4.5. Background variables

The results were discussed above by examining differences between males and females, and taking the geographical factors into consideration, as it is likely that the candidates taking the examination at the three universities largely came from the surrounding geographical districts.

Other background variables that were controlled included such factors as the year of taking the matriculation examination, length of the secondary school course in English, home (city, country town or village), overall grade in the matriculation examination, grade in English, success in the written tests, success in the listening tests, occupation of father, economic background of home, residence in an English-speaking country, and interest in English language outside the school course. The dependence between the candidates' opinions about the appropriateness of entrance tests and these factors was examined using factor analysis (Varimax solution). The results of the analysis are given in Appendix 5. The table shows that the background variables and the opinions about the tests are grouped under different factors, i.e., there is no dependence between them; the existence of such interdependence could be interpreted as a factor that would reduce the validity of the entrance examination. The results of the factor analysis

show, further, that the opinions about the written tests and the listening tests were grouped under different factors, which means that they were assessed in a different way, listening tests having a lower mean and larger standard deviation.

5. OPINIONS ABOUT INTER-UNIVERSITY CO-OPERATION

The 1974 entrance test was planned jointly and administered simultaneously at the universities of Joensuu, Jyväskylä and Turku. With a view to the future extension of co-operation, it would be useful to know the candidates' reactions to such plans. Table 11 gives the results of the two questions enquiring about this.

Table 11. Student opinion about inter-university co-operation and the plans for extending the co-operation, by the sex and the testing place.

Desirability of co-operation and extending it	Joensuu			Jyväskylä			Turku			Total		
	M	F	T	M	F	T	M	F	T	M	F	T
Desirability of co-operation												
Not desirable	6	3	4	0	1	1	15	3	6	5	2	3
Cannot say	19	23	22	17	24	21	20	22	22	18	23	22
Desirable	75	73	74	81	75	76	65	74	73	77	76	75
	100	100	100	100	100	100	100	100	100	100	100	100
The idea of extending co-operation												
Bad	12	3	7	0	1	1	5	7	6	4	3	3
Cannot say	12	17	15	17	17	17	11	9	10	15	13	14
Good	75	80	78	83	82	82	84	84	84	82	84	83
Total %	100	100	100	100	100	100	100	100	100	100	100	100
N	16	30	46	47	144	191	20	90	110	83	264	347

Three quarters of the candidates were in favour of inter-university co-operation, and the idea of extending it to include other universities as well was seconded by four fifths of them; these opinions are independent of sex. Such a positive attitude is understandable, in view of the advantages of co-operation, such as more uniform requirements and selection principles, and the practical and economic considerations.

Less than a fifth were unable to state their opinion here, and the idea was opposed by only 3 per cent of the candidates. The opinions given by the Jyväskylä group were more positive than elsewhere, and the co-operation was received most negatively by males in the Turku group.

6. CONCLUSIONS

The main results of the study can be summarised under the following conclusions.

1. The 1974 joint entrance test may be considered to have been quite successful: it is regarded as difficult rather than easy, and as a better kind of test than mere translation; the weighting of the tests and the school record is found acceptable, and the date of the examination (early July) is suitable to most candidates.

2. The time allocation for the different parts of the examination was sufficient and the test instructions fairly clear; however, the instructions for the reading comprehension test particularly, and general instructions about the examination, need to be improved.

3. The advance information about student selection was not uniform and adequate. More detailed information needs to be given in all the universities.

4. More attention will have to be paid to considerations of fairness. In addition to detailed information about the actual tests, information is also needed about the principles of selection, level of requirement, how to apply for a revision of the selection decision in the case of injustice, etc. The significance of such information is that it will put all candidates to an equal footing, give them a better possibility to evaluate their chances of admission in more realistic terms, and is thus likely to reduce the tension inherent in such an important examination as a university entrance test.

5. The co-operation should be extended to include other universities as well.

REFERENCES:

Carlsson, Gösta and Gesser, Bengt, Universities as Selecting and Socializing Agents: Some Recent Swedish Data. Acta Sociologica 1965, 9, 25-26.

Eskola, Antti, Sosiologian tutkimusmenetelmät II. Porvoo 1968.

Kivekäs, Iris and Parjanen, Matti, Oppilasvalinta ja menestyminen akateemisissa opinnoissa. Tampereen yliopiston opintotoimiston tutkimuksia ja selvityksiä 6/1971.

Komiteamietinto 1971: A 4, Korkeakouluhallinnon oikeusturvakomitean mietintö. Helsinki 1971.

Marin, Marjatta, Tiedekunnat opintoympäristöinä. Helsinki 1970.

Parjanen, Matti, Päällekkäishaku korkeakoulujen oppilasvalinnassa. Tampereen yliopiston opintotoimiston tutkimuksia ja selvityksiä 12/1972.

APPENDIX 1. Advance information for candidates applying to Turku (from Opintielle Turun yliopistoon 1974).

Englantilainen filologia

Valintakoe on pakollinen kaikille pyrkijöille, ja siihen ovat ilman eri kutsua oikeutettuja osallistumaan kaikki hakemuksensa määräaikaan mennessä jättäneet. Koe suoritetaan yhteistyössä Jyväskylän yliopiston ja Joensuun korkeakoulun englannin kielen laitosten kanssa siten, että yhteisesti suunniteltu koe toimeenpannaan samanaikaisesti näissä kolmessa yliopistossa. Kokeen voi suorittaa missä tahansa näistä kolmesta paikasta riippumatta siitä, mihin yliopistoon hakija ensisijaisesti pyrkii.

Hakumenettely: Varsinainen hakemuslomake lähetetään vain siihen yliopistoon, jonka pyrkijä asettaa 1. sijalle (Turku/Jyväskylä/Joensuu: poikkeus: Jyväskylän yliopisto vaatii kuitenkin hakemukset kaikilta Jyväskylän hakemuksessaan jollekin sijalle asettaneilta). Tämän lisäksi jokaisen hakijan on ehdottomasti palautettava tämän oppaan sivulla 46 oleva lomake siihen yliopistoon, jossa hän aikoo suorittaa valintakokeen; tämä on välttämätöntä, jotta tehtäviä voidaan monistaa tarpeellinen määrä. Tämä lomake on palautettava 3.7. mennessä, ja pyrkijä voi siinä halutessaan esittää, että hänen suorituksensa otetaan huomioon myös muissa projektiin osallistuvissa yliopistoissa. Ellei pyrkijän saavuttama pistemäärä riitä 1. sijalle asetettuun yliopistoon pääsemiseksi, tarkistetaan automaattisesti, riittääkö saavutettu pistemäärä 2. tai 3. sijalle asetettuun yliopistoon, mikäli nämä on lomakkeessa mainittu.

Valintapäätöksen ratkaisee valintakokeen tuloksesta ja ylioppilastutkintotodistuksen ja lukion päästötodistuksen alempana mainituista arvosanoista koostuva yhteispistemäärä, jossa koulutodistusten osuus on 30 % ja valintakokeen osuus 70 %. Todistukset pisteytetään edellä s. 30 esitetyllä tavalla ja niistä otetaan huomioon seuraavat arvosanat: 1) yo-todistuksen yleisarvosana (pisteitä 1-9), 2) yo-todistuksen arvosana englannin kielessä ja 3) äidinkielessä (molemmissa asteikko 1-9), 4) lukuaineiden keskiarvo lukion päästötodistuksessa (0-9) sekä 5) koulutodistuksen numero englannin kielessä (4-10). Todistuksista kertyvien pisteiden vaihteluväli on siis 4-46 pistettä.

Valintakokeessa edellytetään oppikoulun pitkän englannin kielen kurssin erittäin hyvää hallitsemista. Tämän toteamiseksi koe käsittää seuraavat osat: 1) monivalintatehtäviä sanaston, idiomien ja 2) kieliopin alalta, 3) luetun tekstin ymmärtämiskoe, monivalintatekniikka, 4) käännöstaidon testaaminen monivalintatehtävien avulla sekä 5) kuullun puheen ymmärtämiskoe, monivalintatekniikka. Koe kestää kokonaisuudessaan 5 tuntia (kirjallinen osa 4 t + kuullun ymmärtäminen 1 t).

Yllä mainittujen osien painokertoimet tulosten tietokoneella suoritettavassa arvostelussa ovat, osien esitysjärjestyksessä, 3 : 1 : 3 : 4 : 1 : 7 : 6 : 5 : 5 (todistukset 30 % + valintakoe 70 %). Painotettu loppupistemäärä tulostetaan asteikolle 0-80. Esimerkki kesän 1973 valinnoista: hyväksymisrajaksi tällä asteikolla muodostui Turun yliopistossa pistemäärä 57, Jyväskylässä 55 ja Joensuussa 51, (näiden kolmen yliopiston yhteiseen valintakokeeseen osallistui 773 pyrkijää, joista kiintiöiden puitteissa voitiin hyväksyä Turkuun 66, Jyväskylään 70 ja Joensuuhun 20).

x x x x x x x

....
Sukunimi Etunimet

Kotiosoite .....

Anon englantilaisen filologian opinto-oikeutta seuraavassa yliopistossa / seuraavissa yliopistoissa, mainitussa järjestyksessä:

1. sija .....

2. sija .....

3. sija .....

(HUOM. Tällä lomakkeella voidaan anoa Turkuun, Jyväskylään tai Joensuuhun tai näihin kaikkiin haluamassanne järjestyksessä).

Suoritan valintakokeen 8.7. ..... yliopistossa.

.....
Allekirjoitus

Täyttäkää tämä lomake ja palauttakaa se viimeistään 3.7.1974 mennessä siihen yliopistoon, jossa aiotte suorittaa valintakokeen. Kuoreen päälle merkitään: Englannin kielen laitos.

APPENDIX 2. Test instruction (read at the beginning of the entrance test).

Englannin valintakoeprojekti
Joensuu - Jyväskylä - Turku
Kokeen suoritusohjeet

Tämän ohjemonisteen saat pitää - kaikki muut paperit on palautettava valvojille!

Kokeessa on kaikkiaan seitsemän osaa: 1. SANASTO ja KIELIOPPI I (71 osiota); 2. SANASTO ja KIELIOPPI II (76 osiota); 3. KÄÄNNÖSANALYYSI (50 osiota); 4. LUETUN YMMÄRTÄMINEN (26 osiota); 5. KUUNTELUKOE, Erottelutesti (50 osiota); 6. KUUNTELUKOE, Kuullun ymmärtäminen (26 osiota) sekä 7. KUUNTELUKOE, Englanninkielinen luento (10 osiota). Näistä osat 1-4 suoritetaan klo 10-14 (Kirjallinen osa) sekä 5-7 klo 15-16 (Kuuntelukoe). Koe on kokonaisuudessaan nidottu samaan monisteeseen, josta kirjallinen osa (valkoinen paperi) käsittää sivut 1-19 ja kuuntelukoe (punainen paperi) sivut 20-27. PUNAISTEN SIVUJEN SELAILEMINEN KIRJALLISEN OSAN AIKANA ON EHDOTTOMASTI KIELLETTYÄ JA JOHTAA VILPPIYRITYKSISSÄ VÄLITTÖMÄSTI KOKEEN KESKEYTTÄMISEEN KO. HENKILÖN KOHDALLA! Väliajan jälkeen jokainen pyrkijä palaa samalle paikalle kuuntelukokeen suorittamista varten - pane siis mieleesi, monennellako penkkirivillä ja millä paikalla istut aamupäivällä. Kun kirjallinen osa on valmis, voit poistua salista ennen määräajan päättymistäkin. Ennen poistumista on käytetyt vastauslomakkeet käännettävä nurinpäin ja tehtävämoniste suljettava. KIRJALLISEN OSAN VASTAUSTEN MUUTTAMINEN VÄLIAJAN JÄLKEEN ON EHDOTTOMASTI KIELLETTY!

Vastauslomakkeiden täyttäminen: Tietokoneella suoritettavaa korjaamista varten kaikki vastaukset on merkittävä oheisille optisen lukijan vastauslomakkeille, joita tarvitaan 5 kullekin pyrkijälle. 1. lomakkeesta on käytettävissä numerot 1-71, 2., 3. ja 4. lomakkeista 1-76 sekä 5. lomakkeesta 1-10; lomakkeiden loppunumerot jätetään tyhjiksi. Lomakkeet numeroidaan tehtävämonisteessa annettujen ohjeiden mukaisesti, ja koetehtävät on numeroitu lomakkeiden numerointia vastaavasti. Huomaa, että vastauslomakkeella on parilliset tehtävänumerot oikealla, parittomat vasemmalla!

Vastausohje: sinun on annetuista vaihtoehdoista valittava mielestäsi oikea, eli se joka parhaiten sopii kussakin tehtävässä annetun lauseen ajatusyhteyteen. Merkitse tarkoittamasi vaihtoehto (A, B, C tai D) kyseisen tentävänumeron ao. kirjaimen alle vetämällä lyijykynällä paksu viiva pilkkuviivojen väliin, esim. harjoitustehtävän oikea vastaus C: A B C D E. Vastaamatta jättäminen ja useamman kuin yhden vaihtoehdon merkitseminen katsotaan virheeksi. Jos haluat muuttaa vastausta, pyyhi edellinen viiva pois ja vedä uusi viiva.

Tarkista, että kirjallisen osan tehtävät ovat selvästi monistettuina tehtäväpaperissasi. Epäselvissä tapauksissa kysy neuvoa kokeen valvojilta. Ehdotus ajankäytöksi: älä kiirehdi turhaan, sillä aikaa on riittävästi; esim. osat 1 ja 2, 2 tuntia; osa 3, 1/2 tuntia; osa 4, 1.5 tuntia. Tulokset 12.7. klo 12 jälkeen.

Ennen kokeen alkua täytetään oheinen henkilötietolomake; kysyttäviä tietoja tarvitaan tutkimustarkoituksiin.

APPENDIX 3 Questionnaire.

Opetusministeriön asettama englannin valintakoeprojekti

Mielipidetutkimus
08.07.1974

Tässä tutkimuksessa selvitellään pyrkijöiden mielipiteitä valintakoeprojektiin liittyvistä seikoista. Pyydän Sinua vastaamaan oman mielipiteesi mukaisesti lomakkeessa esitettyihin kohtiin. Antamasi tiedot käsitellään ehdottoman luottamuksellisesti, eikä kenenkään henkilöllisyys tule missään vaiheessa julki; vastaukset eivät myöskään vaikuta valintapäätöksiin. Jokaisen mielipide on tärkeä, sillä ne vaikuttavat asiasta muodostuvaan yleiskuvaan.

---

koodaus

0 - 3 Nimi: _____ (sukunimi,etunimet)

4 Sukupuoli: 1 mies 2 nainen (ympyröi sopiva vaihtoehto)

5 syntymävuosi: ____ (koodaaja: ____)

6 ylioppilaaksitulovuosi: ____ (koodaaja: ____)

7 englannin koulukurssin pituus: 1 pitkä 2 lyhyt kurssi

8 kotikunta: 1 maalaiskunta 2 kauppala 3 kaupunki

9 koulu, josta tullut ylioppilaaksi, sijaitsee: 1 maalaiskunnassa 2 kauppalassa 3 kaupungissa

10 kokeen suorituspaikka: 1 Joensuu 2 Jyväskylä 3 Turku

11 pyrkiminen 1-sijaisesti: 1 Joensuu 2 Jyväskylä 3 Turku

12 yo-tutkinnon yleisarvosana: 1 a 2 b 3 c1 4 mc1 5 1

13 yo-tutkinnon englanti: 1 a 2 b 3 c1 4 mc1 5 1

14 yo-tutkinnon äidinkieli: 1 a 2 b 3 c1 4 mc1 5 1

15 lukuaineiden keskiarvo: merkitse ympyrällä se numero, jonka ilmoittamaan väliin keskiarvosi jää: 0 alle 7.0 1 7.1-7.3 2 7.4-7.6 3 7.7-7.9 4 8.0-8.2 5 8.3-8.5 6 8.6-8.8 7 8.9-9.1 8 9.2-9.4 9 yli 9.5

16 opettajan antama englannin kielen arvosana: _____

17 koodaaja täyttää kohdat 17 (I osa: ___), 18 (II osa: ___) ja 19 (I + II: ____), jatka vastaamista seur. kohdasta

20 isän/huoltajan ammatti: _____

(merkitse mahdollisimman tarkasti; jos maanviljelijä, tilan pinta-ala: ___ hehtaaria); koodaaja: ____

21 millaisena pidät kotisi varallisuusasemaa? 1 täysin varaton 2 vähävarainen 3 kohtalaisen varakas 4 varakas 5 hyvin varakas

22 oleskelu englanninkielisessä maassa: 0 ei ollenkaan 1 0-1 kk 2 1-2 kk 3 2-3 kk 4 3-12 kk 5 yli 12 kuukautta

23 missä määrin katsot harrastaneesi englannin kieleen liittyviä opintoja koulukurssin ulkopuolella? 0 ei lainkaan 1 jonkin verran 2 runsaasti 3 erittäin runsaasti; mahdollisesti suoritettu ammattitutkinto, johon on kuulunut englannin kielen opiskelua, ja opiskeluajan kesto (esim. dipl.kielenkääntäjä, 3 vuotta): _____

24 Missä määrin katsot, että valintamenettelystä ja kokeesta annettu ennakko-informaatio oli riittävä? 1 aivan riittämätön 2 kohtalainen, mutta ei riittävä 3 jokseenkin riittävä 4 aivan riittävä (koodaus: 0/1)

Jos ennakko-informaatio oli mielestäsi osittainkin riittämätön, merkitse ympyrällä ne kokeen osat, joiden suorituksesta ja/tai luonteesta pyrkijöille annettua informaatiota olisi lisättävä:

(25) 1 Sanasto ja kielioppi I
(26) 2 Sanasto ja kielioppi II
(27) 3 Käännösanalyysi
(28) 4 Luetun ymmärtäminen
(29) 5 Kuuntelukoe, Erottelutesti
(30) 6 Kuuntelukoe, Kuullun ymmärtäminen
(31) 7 Kuuntelukoe, Englanninkielinen luento
(32) (Koodaaja täyttää: summaindeksi kohdista 25-31 _____)

33 Miten selvänä pidit kokeen suoritusohjeita? 1 joiltakin osin epäselvät 2 kohtalaisen selvät 3 kaikilta osin selvät ja ymmärrettävät

Ympyröi ne testit, joiden suoritusohjeita olisi mielestäsi parannettava (koodaus: 0/1)
(34) 1 Sanasto ja kielioppi I (35) 2 Sanasto ja kielioppi II (36) 3 Käännösanalyysi (37) 4 Luetun ymmärtäminen (38) 5 Kuuntelukoe, Erottelutesti (39) 6 Kuuntelukoe, Kuullun ymmärtäminen (40) 7 Kuuntelukoe, Englanninkielinen luento (41) (Koodaaja täyttää: summaindeksi kohdista 34-40: _____)

42 Oliko kirjallisen kokeen suorittamiseen varattu aika mielestäsi riittävä?
1 riittämätön 2 jokseenkin riittävä 3 täysin riittävä

43 Oliko suullisen kokeen suorittamiseen varattu aika riittävä?
1 riittämätön 2 jokseenkin riittävä 3 täysin riittävä

44 Miten tässä kokeessa käytetyt testit ja koe kokonaisuudessaan mielestäsi soveltuvat oppilasvalintojen pohjaksi kielitaitokokeena?
Vastausohje: ympyröi sopivan vaihtoehdon numerotunnus kussakin testissä erikseen; numerotunnukset:
1 = erittäin heikosti 2 = heikosti 3 = vaikea sanoa
4 = hyvin 5 = erittäin hyvin

(44)	Sanasto ja kielioppi I	1	2	3	4	5
(45)	Sanasto ja kielioppi II	1	2	3	4	5
(46)	Käännösanalyysi	1	2	3	4	5
(47)	Luetun ymmärtäminen	1	2	3	4	5
(48)	Kuuntelukoe, Erottelutesti	1	2	3	4	5
(49)	Kuuntelukoe, Kuullun ymmärt.	1	2	3	4	5
(50)	Kuuntelukoe, Engl. luento	1	2	3	4	5
(51)	Koe kokonaisuudessaan	1	2	3	4	5

52 Tässä valintakokeessa on tulosten tietokoneajossa todistusten vaikutusosuus 30 % ja valintakokeen 70 % lopullisista pisteistä, joihin valintapäätös perustuu. Onko tämä suhde mielestäsi hyvä?
1 sopimaton 2 sopii hyvin
(53) Jos näitä prosenttikertoimia olisi mielestäsi muutettava, millaisia kertoimia ehdotat? Koulutodistukset: ___ %, valintakoe: ___ %, yht. 100 % (Koodaaja: valintakoe, ___ %)

54 Olisiko tämäntyyppinen valintakoe mielestäsi parempi kuin kaksipuolinen käännöskoe?
1 käännöskoe parempi 2 en osaa sanoa 3 tämäntyyppinen koe parempi

55 Pidätkö tässä oppilasvalinnassa kokeiltua yhteistyömenettelyä suositeltavana?
1 ei suositeltava 2 en osaa sanoa 3 suositeltava

56 Parhaillaan suunnitellaan yhteistyön laajentamista valtakunnalliseksi oppilasvalinnaksi, mikä englannin kohdalla merkitsisi luultavasti sitä, että koe toimeenpantaisiin samanaikaisesti ja samansisältöisenä eri yliopistoissamme, ja sen voisi suorittaa missä tahansa; opiskelupaikkojen jakaminen tapahtuisi tietokoneella pyrkijöiden ilmoittaman suosituimmuusjärjestyksen mukaisesti (kuten tässä kokeessa). Miten hyvänä pidät tällaista suunnitelmaa?
1 huono 2 en osaa sanoa 3 hyvä

57 Olettaen, että humanistisen tiedekunnan hakuaika tulee päättymään 15.6., mikä seuraavista ajankohdista olisi mielestäsi sopivin valintakokeen pitoajaksi?
1 15.6.-30.6. 2 1.7.-15.7. 3 16.7.-31.7.
4 1.8.-15.8. 5 16.8.-31.8.

58 Miten arvioisit itse menestyneesi tässä kokeessa?
1 0 - 20 % kaikista vastauksista oikein
2 20 - 40 % kaikista vastauksista oikein
3 40 - 60 % kaikista vastauksista oikein
4 60 - 80 % kaikista vastauksista oikein
5 80 -100 % kaikista vastauksista oikein

(59 koodaaja täyttää: menest. _____)

60 Miten vaikeana pidit tätä koetta kokonaisuudessaan?
1 helppo 2 sopiva vaikeusaste 3 vaikea 4 hyvin vaikea

Seuraavassa on varattu tilaa muille mahdollisille mielipiteille ja huomautuksille, joita ei ollut mahdollista esittää yllä olevien kysymysten yhteydessä:

Kiitoksia vaivannäöstäsi lomakkeen täyttämisessä! Antamiasi tietoja tullaan käyttämään kokeen parantamiseen ja oppilasvalintamenettelyn kehittämiseen.

Viljo Kohonen
projektin johtaja

Appendix 4. Advance information of the university of Jyväskylä

(from Jyväskylän yliopiston opinnot ja opiskelijavalinta v. 1974)

Valintakokeet

- englantilainen filologia ma 08.07. klo 9

Ennakkoilmoittautuminen humanistisen tiedekunnan valintakokeisiin on pakollinen. Sen tulee tapahtua viimeistään viikkoa ennen valintakoepäivää ko. aineen laitokselle.

Kullakin hakijalla on oikeus osallistua jokaisen hakemansa aineen valintakokeeseen ilman ennakkovalintaa. Historia-aineissa on yhteinen valintakoe. Hakijan täytyy tulla hyväksytyksi kahdessa valintakokeessa ennen kuin hänet voidaan ottaa tiedekunnan opiskelijaksi. Edellä mainitusta poiketen hakija voi yhteisen valintakokeen perusteella saada opinto-oikeuden joko yhteen tai useampaan aineeseen riippuen siitä, mihin kyseisen valintakokeen aineisiin hän on anonut ja kuinka hyvin hän on menestynyt.

Englantilaisen filologian valintakoe on pakollinen kaikille pyrkijöille (lukuunottamatta ylemmän korkeakoulututkinnon suorittaneita) eikä kokeeseen lähetetä erillisiä kutsuja. Yhtenäisesti suunniteltu valintakoe pannaan toimeen samanaikaisesti Turun ja Jyväskylän yliopistossa sekä Joensuun korkeakoulussa. Jyväskylän yliopistoon hakenut henkilö voi osallistua kokeeseen millä koepaikalla tahansa. Hakijan tulee tiedekunnalle jättämässään hakemuksessa mainita, missä hän aikoo osallistua englantilaisen filologian valintakokeeseen. Lisäksi hakijan on 1.7.1974 mennessä lähetettävä ilmoitus kokeeseen osallistumisesta siihen yliopistoon, jossa hän suorittaa kokeen. Ilmoitus tapahtuu käyttämällä hakulomakkeen mukana hakijoille lähetettävää lipuketta.

Jos hakija haluaa olla mukana valittaessa englannin opiskelijoita kaikkiin kolmeen korkeakouluun tai kahteen niistä, hänen tulee lähettää eri hakemus kaikille asianomaisille tiedekunnille kussakin paikassa määrätyn hakuajan kuluessa. Lisäksi hakijan on asetettava tiedekunnat suosituimmuusjärjestykseen siinä lipukkeessa, jonka hän lähettää ilmoituksena kokeeseen osallistumisesta.

Valintakokeet perustuvat seuraaviin oppimääriin ja edellytyksiin:

- englantilainen filologia: englannin kielen lukion oppimäärä.

Appendix 5. Factor analysis of some background variables and opinions on the appropriateness of the different tests.

Table 12. Rotated (Varimax) factor matrix of the factor analysis of some important background variables (nos. 1-11) and candidates' opinions about the appropriateness of the sub-tests and the whole entrance examination (variables nos. 12-19) as a student selection instrument. This is rather an unusual use of factor analysis, its only aim being to show that the opinions were independent of the background variables.

Variables	Factors I	II	III	IV	V
1. year of matriculation	03	10	-71	11	-20
2. length of school course in English	06	-20	14	09	44
3. location of home (whether city, country town or village)	11	05	15	-09	-49
4. overall grade in matric.exam.	00	79	-27	00	07
5. English in matric.exam.	05	82	-11	-05	09
6. success in written tests	03	76	28	-02	-13
7. success in listening tests	02	64	30	12	-23
8. occupation of father (according to Rauhala)	09	08	-17	-08	64
9. economic background of home	10	-04	-08	20	-70
10. residence in English-speaking country	-09	-03	75	11	-15
11. interest in English	03	26	49	-02	-18
12. opinion on Voc and Gr. I	87	04	-01	02	-08
13. opinion on Voc and Gr. II	86	-00	05	02	-06
14. " Transl. Analysis	65	-01	-14	17	04
15. " Reading Compr.	44	06	01	46	10
16. " Audit. Discrim.	02	-07	-07	80	-04
17. " Listening Compr.	11	-00	02	86	-04
18. " Lecture	14	03	03	75	-01
19. " whole test battery	58	08	03	48	-05
Eigen value	3.4	2.5	1.9	1.6	1.3
%	18.0	13.1	9.9	8.4	6.9

APPENDIX 6

(Item 304):

Twenty-five centuries have passed since the function of the writer was first questioned; it is chastening to see what an unspeakable mess the human mind appears to have made of the answer. Against totalitarians on one side and Surrealists on the other it is equally vain - and, I think, needless - to argue. But even the serious thinkers of all these ages leave us in Babel. One propounds that art is champagne; another, that it is the bread of life; another, that it is poison; another, that it is syrup of figs; another, that it is opium; another, that it is holy water. Between those who damn its consequences and insist on its pleasures, and those who damn its pleasures and insist on its consequences, the truth seems to me to lie, where it so often does lie, simply in the middle.

According to the speaker, the importance of literature is -

a) only in the instruction it gives
b) only in the pleasure it gives
c) in both the pleasure and instruction it gives

(Item 301):

Interviewer: One of your hallmarks is your laugh, isn't it?

Lovejoy: Well, yes.

Interviewer: Have you any interesting anecdotes connected with your laugh? I think you were telling me there was something in Rhodesia -

Lovejoy: Oh, well, there was this - that I never think it's a special laugh at all, but other people seem to think so, and I do remember on one occasion when I was in Rhodesia (after getting back from India we had two years in London and then I went out to Rhodesia for six years), and out in Rhodesia there is a very famous mission run by the Community of the Resurrection called Penhalonga. And we happened to be visiting it on the way to one of the holiday resorts and we just stayed there the night and next day we went on. When I got back to my parish, I was speaking, oh, some months later to someone and they said: 'Were you ever in Penhalonga this summer?' I said: 'Yes. Why?' He said: 'When?' So I gave him the date. He said: 'Isn't that strange? My wife and I were picnicking on the hills above Penhalonga; we could see Penhalonga down in the valley, and suddenly we heard someone laugh. My wife said to me, 'That must be Mr. Lovejoy.' And so it was.

Item 301 The point of this anecdote is that -

a) Penhalonga is a funny place
b) Mr. Lovejoy can easily be recognized by his laugh
c) it was coincidence (=yhteensattuma) that Mr. Lovejoy and his acquaitances were in the Penhalonga area on the same day.

(Item 289):

Ann: Have you heard the news about Betty and Tom?
Bill: No. What about them?
Ann: They're getting married next month. They'll make a wonderful couple, won't they?
Bill: I wonder. What are they supposed to live on?

Item 289

a) Ann thinks that marriage is wonderful
b) Bill does not approve of marriage
c) Bill thinks that people should have a secure income before they get married.

Tiivistelmäkortti

Viljo Kohonen (ed.) (1974) Inter-university student selection for English 1974: a joint entrance test project of the universities of Joensuu, Jyväskylä and Turku. Kielikeskuksen julkaisuja 5. Jyväskylän yliopisto. ISBN 951-677-388-5

Raportissa selvitellään vuoden 1974 englannin kielen oppilasvalintojen yhteistyöprojektia, johon osallistuivat Joensuun korkeakoulu ja Jyväskylän ja Turun yliopistot. Yhteisesti suunniteltu, 309 osiota käsittävä monivalintakoe toimeenpantiin samanaikaisesti näissä korkeakouluissa, ja sen suoritti yhteensä 843 pyrkijää. Hirvosen artikkeli tarkastelee oppilasvalintojen ongelmia ja mittausteknillisiä näkökohtia. Kohonen selvittelee kokeen suunnittelua ja toimeenpanoa sekä esittää tilastollisen testianalyysin päätuloksia. Jarrett tarkastelee käännösanalyysitestiä, ja Hughes ja May analysoivat kuuntelukokeita. Roinisen artikkeli on sosiologinen tutkimus pyrkijäiden ennakkoinformaatiota, käytettyä koetta ja valintamenettelyjen edelleenkehittelyä silmällä pitäen. Tutkimukset on suoritettu opetusministeriön tarkoitusta varten myöntämän määrärahan turvin.

Viljo Kohonen (ed.) (1
selection for Englis
project of the unive
Turku. Reports from
versity of Jyväskylä

The report gives the r
selection entrance tes
taneously at the unive
Turku. The test batte
items and was complete
paper examines problem
etical aspects of lang
brief account of the p
examination and provid
tical analysis of the
lation analysis test,
listening tests. Roini
study of the candidate
procedures and the tes
tions have been offere
student selection. Th
a grant from the Minis

Abstract card

iversity student t entrance test ensuu, JyvÄskylÄ suja 5. JyvÄskylÄn	Viljo Kohonen (ed.) (1974) Inter-university student selection for English 1974: a joint entrance test project of the universities of Joensuu, JyvÄskylÄ and Turku. Reports from the Language Centre No 5. University of JyvÄskylÄ, Finland. ISBN 951-677-388-5
englannin kielen , johon osallis- skylÄn ja Turun 309 oslota kÄsit- anarvoisesti tti yhteensÄ 843 stelee oppilasva- siÄ nÄkÖkohtia. alua ja tulosanta- analyysin nÄÄtu- analyyseistÄ, .koweita. Roini- us pyrkijÖiden ta ja valintame- lÄ pitÄen. Tutki- Ön tarkoitusta	The report gives the results of the 1974 joint student selection entrance test for English, administered simultaneously at the universities of Joensuu, JyvÄskylÄ and Turku. The test battery comprised 309 multiple-choice items and was completed by 843 candidates. Hirvonen's paper examines problems of student selection and theoretical aspects of language testing. Kohonen gives a brief account of the planning and administration of the examination and provides the main results of a statistical analysis of the tests. Jarrett examines the translation analysis test, and Hughes and May analyse the listening tests. Roininen's paper is a sociological study of the candidates' opinions about the selection procedures and the tests. In each paper, some suggestions have been offered for the further development of student selection. The research was carried out under a grant from the Ministry of Education.

Julkaisija:

Valtakunnallinen kielikeskus
Jyväskylän yliopisto
40100 Jyväskylä 10

Published by:

Language Centre
University of Jyväskylä
SF - 40100 Jyväskylä 10, Finland
ISBN 951-677-388-5

83